

CONSTRUCTION OF MORTALITY TABLES FROM THE RECORDS OF INSURED LIVES

RAY D. MURPHY AND PERCY C. H. PAPPS

368.01

Ac8a

v. 2



THE ACTUARIAL SOCIETY OF AMERICA

THE UNIVERSITY
OF ILLINOIS

LIBRARY

~~519.5~~ 368.01

Ac8a

no.2

MATHEMATICS

DEPARTMENT

Return this book on or before the
Latest Date stamped below.


Theft, mutilation, and underlining of books
are reasons for disciplinary action and may
result in dismissal from the University.

University of Illinois Library

MAR 3 1985

MAR 29 1983

DEC 9 REC'D



Digitized by the Internet Archive
in 2017 with funding from
University of Illinois Urbana-Champaign

CONSTRUCTION OF MORTALITY TABLES FROM THE RECORDS OF INSURED LIVES

PRINCIPAL CONTRIBUTORS

RAY D. MURPHY AND PERCY C. H. PAPPS

PUBLISHED BY
THE ACTUARIAL SOCIETY OF AMERICA
256 BROADWAY, NEW YORK

1922

ORIGINAL COMMITTEE

ARTHUR HUNTER

WENDELL M. STRONG

HENRY MOIR

A. A. WELCH

P. C. H. PAPPS

A. B. WOOD

JOHN K. GORE, *Chairman*

COMMITTEE NOW IN CHARGE

ROBERT HENDERSON

WENDELL M. STRONG

J. M. LAIRD

J. S. THOMPSON

A. T. MACLEAN

HUGH H. WOLFENDEN

A. H. MOWBRAY

HENRY MOIR, *Chairman*

COPYRIGHT 1922, BY

THE ACTUARIAL SOCIETY OF AMERICA

NEW YORK

CONSTRUCTION OF MORTALITY TABLES FROM THE RECORDS OF INSURED LIVES.

CHAPTER I. INTRODUCTION.

A mortality table starts with a group of persons at a specified age and shows the number of survivors at each subsequent age. There is also generally set down the number dying in each year of age. The radix of the table, or the number living at the youngest age shown, is arbitrarily selected. Then, from the values of p_x for each age, the values of l_x may be computed successively for the higher ages by the relation, $l_x p_x = l_{x+1}$.*

As a general rule the construction of a mortality table based on the records of insured lives is effected by ascertaining the value of q_x for each age. This value is obtained by dividing θ_x by E_x , where θ_x represents the deaths and E_x the exposed to risk of death for the year between ages x and $x + 1$ in the data under observation.†

In practice some persons will be found who are under observation for only a part of a given year of age either because they enter the experience after the beginning of the year or because they pass out of it for causes other than death before the end of the year. In such cases each person is counted in computing E_x as a fraction equal to the proportion of the year under observation. The degree of accuracy with which such fractions are computed varies according to the method used in tabulating the data. A person who dies between ages x and $x + 1$ must be included in E_x as exposed to risk for the full year and not a fraction, because q_x represents the proportion of l_x persons alive at age x who will not be alive at age $x + 1$.

* See Institute of Actuaries' Text Book, Part II, Chapter I.

† As will be seen in Chapter III, θ_x may be the number of deaths and E_x the number of lives exposed to risk of death, or they may represent respectively the number of policies terminated by death and the number exposed to risk of termination by death; or lastly they may be the amount of insurance terminated by death and the amount exposed to risk. It is important to note, however, that both numerator and denominator in the fraction θ_x/E_x must always relate to the same kind of data.

It should be remembered that theoretically the numerical value of the radix selected and the resulting size of the values of l_x and d_x are of no significance, but that the relative values of these functions are of vital importance. The use of a large radix is advisable, however, so that when the calculated values of l_x and d_x are adjusted to the nearest integer the necessary error introduced is insignificant. This will also have a bearing on the limiting age (ω) since this would be the lowest age for which the unadjusted value of l_x is less than .5, unless fractions are shown.

When an applicant is accepted for insurance after being examined by the company's physicians, he is a "select" life. Among a number of such lives after the lapse of a few years there will be some whose health has become impaired to a greater or less degree, while others will remain as healthy as when first examined for insurance. The survivors of a body of select lives are therefore called "mixed" lives.

It follows that the rate of mortality of insured persons of a given attained age, say x , will not be the same among persons just insured at age x as among persons insured n years ago at age $x - n$. It has been found by experience that persons just insured at age x are subject to a lower rate of mortality than those also aged x , but insured at age $x - 1$. These latter, in turn, usually show a lower rate of mortality than those aged x , but insured at age $x - 2$. In general, for limited values of n , which vary in different experiences and in different age sections of a single experience, it is found that $q_{[x-n-1]+n-1} < q_{[x-n]+n}$ where the portion of the suffix within the square brackets indicates the age at issue, and the other portion, the duration since entry, the total being the present age, *i.e.*, x .

As n increases, the extent of the difference will be found to decrease, so that if $q_{[x-n-1]+n-1} = q_{[x-n]+n} - \delta_n$ then as n increases, δ_n will approach the limit zero.

If t be the greatest value of n for which the relation $q_{[x-n-1]+n-1} < q_{[x-n]+n}$ holds, this fact is expressed by saying that the effects of selection last for t years. Accidental fluctuations in the data, on which the mortality rates are based, are alone sufficient to prevent any exact determination of the value of t . An approximate value is all that can be expected. In practice, t may have a small value, as, for instance, in the case of residents in the tropics, where values as low as 2 years or even 1 year may be found; or it may have such a comparatively large value as 10

years or more, the latter figure applying in the case of the British Offices' Experience (1863-1893) under whole life participating policies, especially at the younger ages. It may be questioned whether the effects of selection ever entirely disappear or whether they become so merged with other influences, such as changes in sanitary conditions and in the mortality of the general population, that they are lost.* As a practical matter, however, we are warranted in assuming that they cease after a certain period.

Now, if a body of select lives all of a given age be observed, and the rate of mortality resulting during the first, second, third, etc., years of insurance be set forth for each year, the result will be a select table of mortality for that particular age at entry. If similar tables be prepared for each age at date of selection, we obtain a set of "select mortality tables."¹

It would involve much labor, however, to base calculations on such a set of select tables. A trial is therefore made to ascertain the effective period of selection beyond which the rate of mortality appears to depend only upon the attained age and may consequently be formed into one "ultimate" table. This may be done by a direct comparison of the values found for $q_{[x-n]+n}$ for different values of n as indicated above. Such a comparison may be confusing, however, because of the large number of values to be observed and the fluctuations in them, and it will usually be more satisfactory to determine by observation approximately where the line of division lies and then apply a different final test.

In describing the construction of the American Men Tables, it was stated that "the crude death rates were deduced for each of the first five insurance years, for the sixth and succeeding years combined, and for the eleventh and succeeding years combined. The expected deaths for each of the first ten insurance years were then calculated by graded rates of mortality based upon the data for the sixth and succeeding insurance years in order that the number of years for which medical selection lasted could be determined. It was seen that the material for the sixth and succeeding insurance years could be safely combined according to attained age."

A different method was used in compiling the $O^{[M]}$ table. The expectations of life were employed, as they would not be subject

* See *T. A. S. A.*, Vol. XIII, page 211, for a discussion of the effect on select tables of a variation in mortality during the period of investigation.

to fluctuations to the same degree as would the mortality rates for individual ages. On page 146 of "Account of Principles and Methods" of that experience are shown, for quinquennial groups of ages, the values of the expectations $e_{[x]}$, $e_{[x-5]+5}$, $e_{[x-10]+10}$, etc., to $e_{[x-25]+25}$. There are also given the values of $e_x^{[5]}$ and $e_x^{[10]}$, the expectations of life found by combining the data for the same attained age, but excluding the data for the first five and first ten years of duration respectively. If selection were still effective in the $(n+1)$ th year, $e_{[x-n]+n}$ would be greater than $e_x^{[n]}$. It was decided for practical advantages to consider that the effect of selection had disappeared in ten years, although this did not appear to be true for all ages at entry.

A select and ultimate table may be set forth conveniently as shown by the following section of the $O^{[NM]}$ table.

Age at Entry.	Years Elapsed Since Date of Insurance.						Age At-tained.
	0	1	2	3	4	5 or more.	
x .	$l_{[x]}$.	$l_{[x]+1}$.	$l_{[x]+2}$.	$l_{[x]+3}$.	$l_{[x]+4}$.	$l_{[x]+5}$.	$x+5$.
20	100,000	99,580	99,003	98,333	97,616	96,879	25
21	99,264	98,844	98,267	97,596	96,877	96,137	26
22	98,530	98,109	97,530	96,857	96,135	95,392	27
23	97,794	97,369	96,790	96,115	95,389	94,641	28
24	97,055	96,630	96,048	95,370	94,639	93,886	29
25	96,316	95,887	95,302	94,619	93,884	93,124	30
26	95,567	95,135	94,547	93,862	93,122	31
27	94,818	94,382	93,791	93,100	32
28	94,059	93,618	93,023	33
29	93,300	92,854	34
30	92,529	35
31	36
32	37
33	38

When the rates of mortality are obtained, the first line of the table may be started with the desired radix and the values successively computed across the first line and then down the last column. The second and subsequent lines may be calculated by working back from the ultimate column by means of the equality $\log l_{[x]+n-1} = \log l_{[x]+n} - \log p_{[x]+n-1}$.

If the data entering into both the select and ultimate sections of the table be combined, or, in other words, if a mortality table be formed according to age only, irrespective of the year of insurance, the result will be an "aggregate table."

While it may be known that the effects of selection last for several years, it may be thought desirable for practical purposes to construct a table of mortality excluding the experience of, say, the first two years only, without constructing the select tables corresponding to those first two years. Such a table is known as a "truncated" table. Every ultimate table is in a sense a truncated table, but the name "ultimate" is usually applied only to a table which forms the continuation of the select section of a mortality table. In the select section the rate of mortality is shown for the age attained, but modified according to the length of time elapsed after initial selection. In an aggregate or an ultimate table the rate of mortality is shown for each age attained, without modification.

It is desirable to consider the effects of the duration of insurance of the data entering into an aggregate table. It will be understood that the following remarks will apply generally but in a modified degree to a truncated table. By first obtaining a clear idea of the nature of these different forms of mortality tables, the student will be in a better position to grasp the significance of the various methods of investigating and collecting the data.

For the sake of illustration, let it be first assumed that at all ages the effects of selection will last for ten years only. Then, if an aggregate table be formed from the experience of a company that has been in business for only ten years, the resulting table will be composed only of lives which have not reached the ultimate rates of mortality. The rates of mortality shown will obviously be much less than will be the case when the table is based upon the total experience of an old company which has been many years in business; for in the latter case, the higher ultimate rates of mortality of the old business will be included, raising the aggregate rates for any given age above the lower mortality of the newer business for the same age. Again, if there be two companies of the same age, the aggregate tables formed from the experience of the respective companies will differ considerably if one company has recently been writing a much larger business in relation to its size than the other. The company having the larger proportion of select business in its aggregate data will, other conditions being the same, show the lower mortality experience. In making this statement it is assumed that the age distribution is similar and that the companies are subject to

the same select rates of mortality, the difference in the aggregate mortality being due solely to the different proportions of new and old business.

From what has been said it will be clear that a select table is the true measure of the adequacy of premiums for life insurance. The great convenience of an aggregate or an ultimate table for actuarial calculations however makes the use of such a table desirable if it can be shown to produce adequate results. The Combined Experience, the H^M and O^M Tables are aggregate tables, the American Experience an ultimate, and $O^{M(5)}$ a truncated aggregate table. These tables have served a very useful purpose as the basis of insurance premiums and reserves.*

For a basis of comparison in mortality investigations select tables are essential to prevent erroneous conclusions due to the difference in the average policy duration of two or more classes of policyholders.

* See *T. A. S. A.*, Vol. XII, page 49, for application of the principle of truncated tables.

CHAPTER II. SOURCES OF DATA.

Any records showing the distribution according to ages or groups of ages of persons living and the number of deaths occurring among such persons, may be used for the construction of a mortality table, provided there are a sufficient number under observation to permit the law of average to operate and to show the general trend of the mortality. The larger the number of lives involved, the more reliable, as a general rule, will be the resulting mortality table. While it is preferable to be able to ascertain the exposed to risk and the deaths for each age, it is not essential; for if the average values of q_x for groups of ages are known, a mortality table may be constructed by interpolation showing the values of q_x for every age.*

Population statistics have been utilized for compiling mortality tables. With the properly compiled records of two censuses and of the intervening deaths a reliable table may be constructed. The Carlisle Table and the English Life Tables are examples of tables which on the whole are properly constructed, while the Northampton Table is an example of an unsuccessful attempt to construct a mortality table from the records of deaths only. The construction of a mortality table from population statistics is covered in No. 3 of Actuarial Studies, and the characteristics of the tables mentioned are set forth in No. 1.

The records of insurance companies and fraternal societies form the most valuable source for compiling mortality tables for the use of such companies and societies. The tables are required for measuring the mortality among practically the same class of lives as that on which the tables are based. The American Experience Table may be cited as one based on the experience of an American life insurance company, namely, The Mutual Life Insurance Company of New York. The American Men Table was formed from the combined data of many American and Canadian companies showing the experience on American males. The Combined Experience, or, as it is sometimes called, the Actuaries', or Seventeen Offices' Table, the H^M Table and the OM Table are based on the experience of groups of

* See Actuarial Studies No. 4.

British companies. The experience of a number of fraternal societies was used as the basis of the National Fraternal Congress Table. These tables are described in No. 1 of Actuarial Studies.

Among the various miscellaneous sources from which mortality tables may be compiled are the volumes, published annually, giving details concerning the families of the British Peerage. These volumes supply data relating to British Peers, the sons and daughters of Peers and the sons and daughters of the eldest sons of Peers, from which a number of Peerage Tables of mortality have been constructed. The records kept by universities have also been employed in compiling a mortality table. From the records of Widows' and Pension Funds tables of mortality may be formed which will be useful as a guide to the management of such funds.

CHAPTER III. METHODS OF RECORDING THE DATA.

As this study deals only with mortality tables from the records of insured lives the methods hereafter described will be those applicable to the data available in an insurance company or fraternal society. The facts regarding each policy will be set down on a card so that they may be sorted and tabulated (See Chapter IV). The details to be recorded will depend partly upon the use for which the mortality table is desired, and methods may be varied to facilitate the work, provided accuracy is not thereby sacrificed.

There are three possible bases for determining q_x , (1) the death rates among the LIVES insured, (2) the rates of termination by death of POLICIES in force, and (3) the rates of termination by death of the AMOUNTS insured.

If it is desired to investigate by lives, the data concerning all the policies on each life must be brought together to avoid duplication. Where the experience is confined to one company the office records will usually show all policies on a single life, but when the combined experience of several offices is under investigation, the existence of duplicate policies in different offices can be ascertained only by examination. The usual plan is to arrange all names in strict alphabetical order—as in a directory—and thus enable the bulk of the duplicates to be brought together. Where names agree, but the dates of birth differ radically, different lives may be presumed; where the dates of birth differ slightly, it may sometimes be found that the cards refer to a single life. Occasionally day and month of birth will agree, but not the year. Inquiry as to occupation and residence will usually throw light upon doubtful cases.

Some persons, however, are careless in the use of their names, and it will occasionally happen that a name may appear as, say, George Frederick Smith in the records of one company, and Frederick Smith in those of another—or the surname may be written Smyth. Alphabetical sorting will fail to disclose the existence of such duplicates, and, where complete elimination is considered important, an independent sorting of the material by dates of birth should be effected. This latter device is

especially useful in the case of female lives where change of name owing to marriage is quite often recorded in one office and not in another, owing to carelessness on the part of the insured, or to the termination of a policy prior to marriage. Even so, errors in dates of birth will prevent complete elimination of duplicates.

Having brought the duplicates together the method of elimination depends upon the kind of table that is to be constructed. If it is an aggregate table and there has been a continuous exposure from the date of entry of the earliest policy to the last date of exit, one card may be written for the complete exposure and the others destroyed. If there has not been a continuous exposure then two or more cards must be written to cover the several continuous periods that the company was at risk. If select tables are to be compiled the life should count only once in the ultimate section of the table, but may be counted once for each select period beginning at a different age. This is seen from the fact that each such select period is a separate experience and it involves no duplication to allow one life to enter two or more such experiences.

A difficulty enters here because the duration of the select period cannot be determined in advance. One card may therefore be written for the complete exposure and separate cards for the other select periods, using an assumed duration for the period of selection. These latter cards should be coded so that they may subsequently be sorted out and the durations corrected after the select period has been decided upon.

The British Offices' Experience illustrates the practical difficulties which may arise in eliminating duplicates. Several sections of the experience were to be compiled independently, combining the data in some of these afterward. Furthermore aggregate, truncated and select tables had to be provided for. Strict elimination in every case would have been too laborious and consequently duplication was considered in each section only. When sections were combined some lives were therefore recorded more than once. Furthermore one card could not well be written for the continuous exposure on one life for the aggregate table, as this would conceal the select experience arising from policies issued after the earliest one which should properly enter the select tables, and also would make the truncated tabulation desired impossible. A card was consequently

written for each age at which a policy was issued but such cards were marked to enter the aggregate experience only at the duration corresponding to the date of exit of the preceding policy. The result was that the select tables contained duplication in the ultimate section arising from policies on one life issued at different ages. The truncated table likewise contained two or more coincident exposures on some lives. The investigation was therefore not based wholly on lives but on a combination of lives and policies.*

Mr. G. F. Hardy has stated, in "The Theory of the Construction of Tables of Mortality, etc.," p. 18:

"Without dogmatizing upon the point, it appears to me that the proper course is, where two or more policies are effected at the same time or at the same age at entry, to treat them as a single risk, but where the subsequent policies are effected at later ages, involving fresh medical selection, to treat them as separate risks. This means the elimination of duplicates in each of the 'select' tables for individual ages at entry, but no further elimination in the resulting aggregate tables, a course which has the advantage of making the aggregate table the true aggregate of the tables for separate ages at entry. Judging by the results of the O^M experience, this course is necessary if we are to produce an aggregate table, representing 'ultimate' rates of mortality after the lapse of a stated period from entry, which will join on smoothly to the 'select' rates."

If the investigation is by policies there will be cases where one life is insured under several policies and the failure of one such life will have a greater effect upon the mortality table than the failure of a life insured under a single policy. Now, if the lives which are insured under several policies are on the whole subject to a lower rate of mortality than lives of similar ages insured under single policies, it follows that the table based on policies will show lower rates of mortality than if the table had been based on lives and vice versa. The general effect of investigating by policies is to produce rates of mortality fairly close to those determined by lives insured except for accidental fluctuations where the exposures are small; but it has been thought that an experience by policies may show a slight tendency to diminish the values of q_x at the younger and increase them at the older ages in an aggregate table. This is probably due to the fact that at the younger ages, where some individuals are repeatedly

* For an instructive detailed account of the manner in which the cards were marked and duplicates eliminated see the volume of the British Offices' Life Tables, 1893, entitled "Account of Principles and Methods."

undergoing medical examination to obtain additional policies, those who are successful represent a class of superior lives, and the additional weight given to them in the experience tends to lower the mortality; while those who are unsuccessful do not show a proportionate effect. At the older ages where few additional policies are taken this extra weight operates to increase the proportion that the ultimate data bears to the whole.

The saving of labor in investigating by policies rather than by lives will frequently justify the former method. The result will usually be a satisfactory basis for comparison with other mortality tables and with an insurance company's experience in special classes of risks. Neither of these methods however is necessarily a safe basis for the computation of premiums and reserves, since the rate of financial loss may be greater than the rate of mortality by lives or policies. This will be the case when the mortality among lives insured for large amounts is greater than among those insured for small and moderate amounts.

Investigations by amounts insured, in which q_x is determined as the ratio of claims incurred to insurance in force, are considered essential under such circumstances, which are common in this country. It is evident that mortality rates so derived will be subject to accidental fluctuations caused by the failure of one or more lives insured for large amounts. Such fluctuations may subsequently be removed by graduation or may be lessened by restricting the limit of insurance for which any one life will be counted as exposed in the experience. This latter restriction would necessitate the labor of bringing together for investigation all policies on any life insured for a large amount. The use of amounts insured introduces additional work also because of the necessity for noting all changes in amount after issue.

In the American-Canadian Mortality Investigation (1900-1915) it was decided to count as \$100,000 only, insurance on any individual life issued at any one age for more than that sum. In order to accomplish this purpose there was noted on the cards furnished by each company the insured's initials and full date of birth if the policy was for \$50,000 or more. For each age at entry those cards were brought together which had the same date of birth and initials. The mortality was to be investigated also by plans of insurance and Volume I of that experience

states, "When more than \$100,000 of insurance was issued at the same entry age on the same life but on different plans of insurance, the amount of insurance on each card was reduced proportionately, provided the mode of termination and duration were identical. For instance, if \$50,000 had been issued on the twenty-payment life plan and \$100,000 on the ordinary life plan, these amounts were reduced to \$33,300 and \$66,700 respectively. This procedure made it unnecessary to make any further adjustment when the investigation was made by plan of insurance.

"When the duration of the several policies differed, the policy with the longest duration was retained in the investigation for its original amount, provided such amount did not exceed \$100,000."

This method guarded against an undue effect during the select period of the failure of a life insured for a very large amount. There was not a similar necessity for restricting the amount in the ultimate portion of the tables because of the larger volume of exposures during that period. It will also be noticed that the limit of \$100,000 might be exceeded even during the select period through the issuance of policies for less than \$50,000 at the same age, but this was not a serious practical objection to the method followed.

In this investigation it was decided to terminate the exposure if the amount insured under a policy was increased or decreased. This eliminated much labor and avoided complications regarding the select period where such a change was made upon evidence of insurability.

A life may pass out of observation by any one of four modes of termination: (1) Existing; (2) Withdrawn; (3) Matured; (4) Died. The exposure of a life terminates by "existing" when the policy is in existence at the close of the period covered by the experience. Lapses and surrenders are classified as withdrawn. All involuntary withdrawals such as expired term policies and matured endowments are treated as matured. The withdrawn and matured have often been grouped together, while on the other hand additional classifications may be adopted depending on whether the experience is to be used for additional information, such as the rate of voluntary withdrawal.

In recording the exposures and deaths in a mortality experience three methods are available; the "Policy Year," "Calendar

Year," and "Life Year" methods. These terms apply to the manner of analyzing the exposures and deaths.

A. Policy Year Method.

Under the policy year method the exposures are traced from the beginning to the end of each policy year and each death is allocated within the exact policy year in which it occurs. The age at entry may be taken as the age nearest birthday, or it may be taken as the mean age, found by subtracting the calendar year of birth from the calendar year of issue. While in any particular case the mean age may be nearly one year greater or less than the correct age, these discrepancies may be considered to balance. In the United States the nearest age method can be followed by extracting the age directly from the policy records. In Great Britain and Canada, however, it is customary to insure at the age next birthday and either method would require a calculation.

As a policy year investigation ordinarily begins and ends with policy anniversaries in specified calendar years the existing will usually pass out of observation at an integral age determined by adding to the age at issue the duration found by subtracting the year of issue from the year with which the experience closes.

In treating the withdrawn fractional durations are involved. These may be treated as the "exact," "nearest," or "mean" duration. The exact method is followed by tabulating the precise fractions of a year of exposure over integral years (in practice usually to the nearest month). The nearest duration is found by taking the nearest integral number of years, with proper adjustment so that in cases where the fraction is $\frac{1}{2}$ the number of cases counted as the next higher integral year will balance those counted as the next lower. The mean duration is found by subtracting the year of entry from the year of exit on the supposition that the overstated durations will approximately balance those understated. The exact method is usually laborious and one of the other methods is therefore to be preferred if it can be found to give approximately correct results. The nearest duration method may in some instances materially understate the exposures; for example, in the first policy year where the lapses at the end of three months may overbalance the lapses at the end of nine months and thus affect the exposures for the first policy year. The mean duration method is

based on the assumption of an even distribution of business throughout the calendar year, which may or may not be sufficiently in accordance with the facts in any particular investigation. In general it may be remarked that neither the nearest nor the mean duration method should be adopted without taking into consideration all the peculiarities of the fundamental data.

Those cases which pass out of observation through maturity seldom introduce fractional durations, but if they arise fractions may be treated in the same way as for the withdrawn.

Deaths must always be treated with accuracy as any error would be of vast importance compared with a similar error in the exposures. As a life is treated as exposed to risk during the whole year of death the duration would naturally be taken to the policy anniversary following death. The curtate (next lower integral) duration, however, may be recorded provided that, in tabulating the results, the exposures are adjusted to include the year of death.

As the policy year method tabulates the experience in the form which is desired for the construction of select tables it is the method which would naturally be employed in important investigations of that character. It was used as the basis of the British Offices' Experience (1863-1893), the Medico-Actuarial Investigation (1885-1909), and the American-Canadian Mortality Investigation (1900-1915). A full description of these applications of the method may be found in the published volumes of those experiences.

B. Calendar Year Method.

The calendar year method is an important historical method, used first by Mr. Woolhouse in the 17 Offices' Experience, and used later in the Institute (20 Offices') Experience and the Thirty American Offices' Experience. It should be noted that this method can be used where the policy details are not sufficiently complete for a policy year method. In connection with its use in the Institute Experience it was stated that much valuable data could not have been included in the experience had precise dates of birth, entry and exit been asked for. All that was available in many cases was the office entry age next birthday and the calendar years of entry and exit. Policy year tabulation was therefore impossible.

The method has however lost favor as the basis of important mortality tables constructed from life insurance experience because it does not produce data in convenient form for select mortality tables. The material may be separated by calendar years after entry and select tables compiled by an approximate method as hereafter described; but year 0—the calendar year of entry—represents only approximately the first six months experience and year 1 extends approximately from duration $\frac{1}{2}$ to $1\frac{1}{2}$, etc. Therefore, if reliable select tables are desired the policy year method should be used. It should also be noted that, under the assumption that those entering and withdrawing do so in the middle of the year, persons withdrawing in the calendar year of entry are excluded from the experience, thus tending to understate the exposures.

Under the calendar year method of analyzing exposures and deaths the calculations in connection with the exposures are usually based upon the assumption of an even distribution of issues throughout the year. Thus new entrants may be assumed to be exposed on the average for six months in the year in which the policies are issued and similarly withdrawals can be considered as taking place in the middle of the year of exit.

The new entrant is assumed to enter in the middle of the calendar year, which is the period used as the basis for grouping the exposures. In this first calendar year therefore he will be treated as exposed to risk for one half a year at his age at the beginning of such calendar year. He will be counted as exposed for the next calendar year at his age at the beginning of that year. Therefore, if the age at date of entry were calculated as the age nearest birthday or mean age at that date, he would enter as a fractional exposure for that year at an age $\frac{1}{2}$ year less and would be exposed throughout the next year at an age $\frac{1}{2}$ year more. So if the age at date of entry were calculated as an integer, x , the calendar year exposures would be at the fractional ages $x - \frac{1}{2}$, $x + \frac{1}{2}$, $x + \frac{3}{2}$, etc. The deaths would also have to be tabulated by half ages. The completed age at death would be the half age on the January 1st preceding death. We may find the completed age by adding the proper duration to the age at entry. If death occurs in the first calendar year following the year of entry we should add $\frac{1}{2}$ to that age, if death occurs during the second year after the year of entry we should add $1\frac{1}{2}$, etc. Thus the general rule for determining the completed age would

be to add to the age at entry (nearest birthday or mean age) the mean duration minus one-half. For example, if an insured were 35 at the date of issue in 1890 and died in 1912, his age on January 1, 1912 would be $35 + (1912 - 1890) - \frac{1}{2}$, or $56\frac{1}{2}$. In this way the exposures and deaths would be grouped by half ages and from the values of $q_{x+\frac{1}{2}}$ so obtained the values of q_x might be interpolated.

It would seem preferable in many cases to derive the rates of mortality directly for integral ages, but in order to accomplish this we must follow a method which will assume that integral ages among the exposed to risk coincide with January 1st. Where the practice of insuring lives at the age next birthday exists, it is convenient to consider that such ages will be attained on the January 1st following entry. This is equivalent to the assumption that the age next birthday is on the average $\frac{1}{2}$ year greater than the exact age at date of entry and that on the average six months elapses between the date of entry and the beginning of the next calendar year. Each life will then be exposed to risk for $\frac{1}{2}$ year in the calendar year of entry at an integral age one year less than the age next birthday at issue, and the succeeding years of exposure will obviously be at integral ages also. Withdrawals would be exposed for $\frac{1}{2}$ year at the integral age corresponding to January 1st of the calendar year of withdrawal.*

Another method that will group the data by integral ages consists of using as the age on January 1st following issue the age nearest birthday on January 1st. If the age so calculated is x , the life would be exposed for half a year at age $x - 1$ in the calendar year of entry.

If the exposures were to be tabulated for integral ages by calculating the age nearest birthday on January 1st following the date of entry or by using as the age on such January 1st the age next birthday at entry, then the age on January 1st of the year of death may be taken by adding to such an age the mean duration minus one. This gives precisely the same result in the former case as would be obtained by calculating the age

* See Mr. Geo. King's remarks, *J. I. A.*, Vol. XXVII (pg. 218) for alternate method of treating new entrants and withdrawals. It should be remembered that the fractional exposures in the years of entry and withdrawal might be calculated accurately but such refinement would destroy the convenience of the method.

nearest birthday at the beginning of the year of death. In the latter case the deaths would ordinarily be given somewhat more accurately by calculating the age nearest birthday on the January 1st preceding death instead of employing the rule just given, which uses the age next birthday at issue. It must be remembered that whatever method of determining the age is used, each death must be included as an exposure for the entire calendar year of death.

For the withdrawn, the age at the beginning of the year of withdrawal may be determined in accordance with the considerations just mentioned. If the age next birthday at issue was used, however, it would be satisfactory to add the mean duration minus one rather than to calculate the age nearest birthday on January 1st of the year of withdrawal as is sometimes done for deaths.

C. Life Year Method.

The third way of analyzing the exposures and deaths is the life year method. This might be called the exact method as it traces the data from birthday to birthday. The deaths are grouped according to the exact age last birthday at the date of death. The exact ages (to nearest month, or to one decimal place) at entry and exit of all lives entering into the experience may be calculated, each life being ordinarily exposed for a fraction of a year of age at the beginning and end of its exposure. Instead, however, of calculating each of these fractions with exactness an average fraction may be assumed, after a sufficient test, to apply to all cases except deaths, which are treated as exposed to risk for the whole year of death.

In experiences where it is justified a rougher approximation may be adopted for computing the exposed. This consists in taking the ages at entry and withdrawal as the mean ages found by subtracting the calendar year of birth from the calendar year of entry and exit respectively. This avoids fractional durations and is based on the assumption that the resulting over-exposures and under-exposures will balance. There is a difficulty presented if the experience starts at a certain date with a number of lives then under observation or closes with a number existing. For such cases the ages at entrance and exit may be taken as the age nearest birthday or determined by some other suitable method.

Another method of estimating the exposures is available by using the ages nearest birthday at entry and exit for the entire data except deaths. Under this method the error involved in the age at entry or exit for any one life cannot exceed six months, while the mean age may be incorrect by almost a year; but under both methods the largest possible error in duration is approximately one year.

The life year method, while making possible the strictest accuracy if the ages are exactly calculated, has a very narrow scope of usefulness in life insurance experience, because it does not produce the results in the form of select tables and, if applied with exactness, is laborious. It has been employed in tabulating mortality among annuitants.*

A general warning should be given regarding the use of approximations which have been frequently suggested in connection with all these methods. In any important investigation the effects of such approximations should be tested before they are adopted. For example, where an office accepts applicants at rates for the nearest birthday there will be a tendency for persons to insure just before the nearest birthday changes. In such case the nearest birthday at entry would on the average somewhat understate the true age.† If lives are accepted at rates for the age next birthday the tendency will be to insure more often just before than just after a birthday, and the result of using the nearest age may be to overstate the age slightly. Again a company's business may not be evenly distributed over the year, because of pressure at the end of its financial year or toward the culmination of agency contests, and methods depending upon such even distribution may produce too great an error. Considerations such as these should be taken into account in determining the methods to use.

Mortality investigations frequently cover the experience under all policies in force within a specified period. It then happens that many lives will enter the experience at the beginning of the period after having been insured for some time.

* For applications of the method see *J. I. A.*, Vol. XXXI, pages 171-5 and 179-181.

† If a mortality experience is to be used as a basis for premium rates charged at nearest ages there is no objection to such understatement of ages in the data with resulting slight increase in mortality rates; for the companies will thus have provision in the mortality table for the effect of charging premiums at nearest age.

If the policy year method is to be used the experience is generally taken between the policy anniversaries in the first and last years of the period. By this means fractional exposures are avoided at the beginning in the case of lives brought forward. Care must be taken however to see that the policy years elapsed since original entry are properly recorded for use in tabulating select tables. If other methods are used the manner of handling the lives that enter the experience in this way must be consistent with that employed for lives first insured during the period of investigation.

It is customary to allow policyholders a grace of thirty days or one month for the payment of renewal premiums. If death occurs during the grace period while the premium is unpaid the claim is allowed, the overdue premium being deducted from the face of the policy. As such deaths will be recorded in a mortality experience theoretical accuracy might seem to require treating every lapsed policy as exposed to risk for one month after the due date of the premium, in order that the corresponding exposures may be recorded; or, if these fractional exposures are not taken into account, then it might appear that no account should be taken of the corresponding deaths.

It is impossible, however, to ascertain what deaths should properly be eliminated under the latter alternative because we ought to exclude only those cases where the premiums would not have been paid had the person insured survived the grace period. Furthermore there are practical objections to treating the grace period on lapsed policies as exposures, because companies will continue to grant a grace privilege and their premiums must be sufficient to provide that benefit. By including the deaths and excluding these fractional exposures the mortality table will give premiums which contain a provision for this. It is therefore the usual practice to include the deaths and exclude the exposures.*

When a policy lapses after having been carried a short time, it is customary to grant an automatic nonforfeiture feature. If automatic extended insurance is given then it must be decided whether the exposure is to cover the period of risk under the original form of policy only, or whether the risk under the

* The British Offices' Experience is an important exception to this rule. The nearest duration method was modified to take account of the grace period. See Appendix M, volume entitled "Account of Principles and Methods."

term extension is to be included. When the decision is made the deaths are taken to correspond. If the risk under the extended term insurance is to be included the fractional durations at the expiry of the term insurance must be treated in the same way as the fractional exposures of the withdrawals.

Many policies are changed at lapse to reduced paid-up insurance either automatically or upon request. If the experience is based upon amounts insured, such policies, if included in the experience, must be treated as withdrawals for the amount of the reduction and the balance continued. This is awkward and it may be thought best to exclude all data from the date of lapse.

It must also be kept in mind in treating paid-up and extended insurance that the character of the mortality under these options may vary considerably from that under policies in full force. It may therefore be best in some circumstances to make separate investigations for such lapsed policies to determine the company's attitude toward the non-forfeiture features and to exclude the data from the company's general experience.

The question also arises whether to include policies after the amount of insurance has been changed. Except where the investigation is by amounts insured a reduction in amount presents no difficulties, but as an increased insurance is usually granted only upon medical examination such lives would be select from the date of change. If an aggregate table is to be constructed this might be disregarded but for a select table recognition must be taken of this fact. Because of such complications it will sometimes be best to exclude one or both of these classes of changed policies.

The same principles will control the decision as to the inclusion of policies changed in form, for changes to lower premium forms are made subject to examination.

If a policy lapses and is afterwards reinstated upon evidence of good health it must be decided whether the exposure should terminate at the date of lapse and start again at the date of reinstatement, or whether the temporary lapse should be disregarded. Where the period between the date of lapse and date of reinstatement is short and only a personal certificate of health has been furnished, no appreciable error will be introduced by ignoring the lapse and reinstatement. Where considerable time elapses over which the policy is not exposed to risk and a new medical examination is made, it may be thought better to take

account of the lapse and terminate the exposure at that point. The subsequent reinstatement may then be treated as a new policy; but, bearing in mind that a company is not always so particular in reinstating a lapsed policy as in issuing a new one, and often reinstates an impaired life covered by extended insurance because the mortality loss will be less under the premium-paying policy, it may be thought better to eliminate entirely from the investigation the exposures under such reinstated policies. It is not likely that any material effect upon the results of the investigation would be caused either by including or excluding these exposures.

In the construction of mortality tables many points will arise which call for the independent judgment of the compiler, and when the tables are published it is usual to state how these different points have been handled.

When constructing such tables it is sometimes desired to investigate collateral questions such as the relative mortality by plan of insurance, amount of insurance, residence, occupation or other features of the risk. It is therefore necessary to plan the card on which the policy details are to be recorded so that provision will be made for the information necessary for such investigations.

CHAPTER IV. HANDLING THE DATA.

The compilation of a mortality experience is usually carried out, as has been mentioned, by recording the data on cards because of the ease of sorting the material by age at issue, duration, mode of exit or any other feature. It is necessary to exercise care to see that the sources from which these details are drawn are reliable, particularly in regard to the mode of exit; for it is essential that the deaths be recorded correctly. It may be worth while to check all cases recorded as dead with the claim records, if they are available. If the cause of death is to be used for a subsequent investigation of the relative frequency of the different causes, such a check may be obtained with little or no extra labor.

If the table to be constructed is an aggregate table or simply excludes certain early years of duration, as in the case of a truncated table, it is necessary to record the age at entrance into the experience and at exit. Equivalent facts would be obtained by recording the age at entry and duration, from which the age at exit can be obtained by addition. When select tables are desired it is necessary to show the duration instead of the age at exit.

In writing the cards the policy records must be examined carefully to exclude the data which is not desired, as for example, under-average lives, female lives, policies continued under non-forfeiture provisions, residents outside the temperate zone, etc.

If the mortality table is to be based on lives insured, the date of birth and name, or all initials, should be entered so that duplicates may be eliminated as described in Chapter III. For a table based on amounts insured care must be taken to record properly cases in which a change in the amount of insurance has occurred. If such cases are to be included in the experience after the date of change it will probably be best to write two cards, one for the lesser amount and one for the balance, with proper dates of entrance and withdrawal to make them equivalent to the case as a whole. When an experience both by policies and by amounts is desired, the lesser amount may be counted as one policy while the balance may be indicated as "no policy."

Then in combination they will count as one policy for the full amount.

For select tables cases re-examined medically, such as increased amounts, changes to lower premium forms, and possibly reinstatements, if included in the investigation, would usually be recorded as withdrawn at the time of change or lapse, and additional cards would be written to tabulate them as new entrants after such dates.

Policies which enter the experience at the beginning of the period covered by the investigation, having been previously issued, must be specially treated. If only an aggregate table is desired the age at entry may be recorded as the age at date of entering the period. If, however, select tables are to be constructed the card must show the age at issue, duration at entry into the experience, and duration at exit.

Those policies which are still in effect at the close of the period of investigation will have the duration at that time recorded.

The form of card to be used should be selected with care to include all information necessary to permit of the easiest and most accurate recording, and to make the elimination of duplicates, sorting and tabulating as rapid as possible. The volume entitled "Account of Principles and Methods" of the British Offices' Experience (1863-1893) gives an excellent example of the points which must be considered. The card shown on the following page is one of several forms used in that experience.

The letters S.A. indicate that this card was to be included in the data both for the select and also for the aggregate tables. The age at entry was taken as the age nearest birthday. Therefore + was marked after the year of birth in this individual case to indicate that the age at entry was the difference between the year of entry and one more than the year of birth. The data were analyzed by policy years, and the duration recorded on the cards for deaths was the curtate duration, or the duration at the commencement of the policy year of death. Cards of different color were used for male and female lives.

It is interesting to note that a preliminary test showed that a general alphabetical arrangement of the cards would not satisfactorily eliminate duplicates. The cards were therefore sorted chronologically by date of birth. After noting the duplicates detected in this way, the cards *in each year of birth* were sorted alphabetically and again examined for duplicates.

S.A.		Old Policies.	
NO.....2356.....	£.....100.....		
CLASS.....0.....	PROFIT OR NOT...P.....		
LIFE	{ Latimer..... { Darsie.....		
<hr/>			
DATE—	D.	M.	Year.
OF BIRTH.....	7.....	8.....	1826..... +
OF ENTRY.....	1.....	1.....	1850.....
In 1863.....			1863.....
OF EXIT.....	29.....	2.....	1888.....
<hr/>			
Duration before 1863.....			13.....
Duration of Policy.....			38.....
<hr/>			
Age at Entry.....			23.....
Age in 1863.....			
Age at Exit.....			
<hr/>			
MODE OF EXIT.....			(D).....
<hr/>			
REMARKS.			
<hr/>			
L			

(Select and Aggregate Tables.)

In compiling the census returns of 1890, the United States Census Office made the first use of the system of punched cards, which was devised by Dr. Hollerith and permits of mechanical sorting and tabulating. This system makes use of a card printed with a series of columns running from 0 to 9. If the card is to represent a policy issued at age 35, the "3" in the tens column and the "5" in the units column representing the age will be punched out. By means of code numbers the occupa-

tion or other information may be entered upon the Hollerith cards; for example, two columns will cover 100 and three columns 1,000 different occupations. By an ingenious mechanical device the capacity of the cards may be somewhat increased beyond the 100 or the 1,000 without increasing the number of the columns beyond the two or three. After the cards are punched they are sorted mechanically at a high rate of speed, it being claimed that 15,000 cards may be sorted in one hour. Another machine specially designed will add four items on the cards at a speed slightly lower than that of the sorting machine. Machines may be built on the same lines to add different numbers of items. The student is referred to two articles (*T. A. S. A.*, XI, 252, 276) for further information concerning the Hollerith system.

Another system which employs similar punched cards is that put out by the Powers Accounting Machine Company. These machines were used in carrying out the American-Canadian Mortality Investigation. A copy of the card used in that experience is given on page 27. (See Page 11, *A. C. M. I.*)

Volume 1 of the investigation states "The machines used for perforating the cards were so constructed that the cards automatically came into place for perforation. The hole to be punched in each column of the cards is selected by a sliding scale or bar and the perforation is done by electricity. Two cards may be prepared at one time, and accordingly, cards of two colors are inserted alternately in the package of blank cards and placed in the machine. The verification of the accuracy in perforating the cards may be done in two ways, either by comparing the perforated card with the original record, or by using a verifying machine so constructed that when the operator strikes an incorrect key or when the key struck does not agree with the hole already in the card the machine automatically stops and thus attention is called to an error on the part of the original perforating clerk or of the operator of the verifying machine.

"A great advantage of this perforating machine is that in certain types of perforation the keys need not be changed if the same figures appear in a number of cards. If, for example, the cards were all issues of the year 1914 and were issued in the State of New York on the twenty year endowment insurance plan, then in the proper fields the figures 14, 3 and 38 would remain automatically in place until all the cards coming within that group had been perforated. Another advantage is that the complete setting of the machine for perforation can be read along the edge of the machine before touching the lever which perforates the card."

CARD USED FOR RECORDING THE DATA.

[illegible]

The following statements regarding the sorting and tabulating machines will give a good idea of the remaining operations of this system.

"The sorting machine if kept supplied with cards, will classify 12,000 per hour. This, of course, refers to sorting any one column. If the cards are to be sorted by age, two sortings are necessary,—first, into groups of decennial ages, *i.e.*, 20–29, 30–39, etc.; and then into separate age groups. Advantages of this type of machine are that the cards are placed by machine in compartments which are easy to empty, as they do not involve stooping, and that all cards of a specified number can be segregated, leaving the cards with other numbers in the same numerical or other order. Several machines used by the bureau had counters, technically known as counting sorters, attached to the sorting machines. These machines count the number of cards going into each compartment or pocket and give also the total number of cards in all the pockets.

"This machine (the tabulator) prints on strips of paper such fields in the cards as are desired. This is known as selecting mechanically the desired columns of the cards to be recorded on the list. It can be so arranged as to record letters instead of the figures appearing on the card. For example, the code number for the termination of the policy by death is 9, but it appears on the list prepared by the machine as 'D.' From one to seventeen columns may be added. The work of this machine can best be understood by means of the following example:

EXAMPLE OF WORK BY TABULATOR-PRINTER.

Explanation of Columns of Figures or Symbols.

- (a) Year of issue; *i.e.*, 2 stands for 1902, 12 for 1912 and 8 for 1908.
- (b) Age at entry.
- (c) Duration in years.
- (d) Sex (M = Men).
- (e) Mode of termination (L = Lapsed).
- (f) Code number for plan of insurance.
- (g) Code number for habitat at date of application for insurance.
- (h) Number of policies.
- (j) Amount of insurance to nearest \$100; *e.g.*, 10 = \$1,000, and 25 = \$2,500

(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(j)
2	20	1	M	L	1	16	1	30
12	20	1	M	L	1	16	1	10
12	20	1	M	L	2	16	1	150
13	20	1	M	L	1	16	1	50
11	20	1	M	L	3	16	1	100
8	20	1	M	L	4	16	1	20
10	20	1	M	L	1	16	1	10
10	20	1	M	L	1	16	1	30
1	20	1	M	L	1	16	1	10
6	20	1	M	L	2	16	1	10
11	20	1	M	L	1	16	1	10
11	20	1	M	L	3	16	1	300

(Total number of policies) 12 (for ins. of) 730
[\$73,000]

"The parts in brackets are not printed by the machine but are added for the sake of clearness.

"After the cards have been sorted into the desired groups, special cards known as 'total' cards were placed between these groups. The machine continues to print as shown above until it comes to a 'total' card when the desired summations are made; the machine then commences to tabulate the ensuing group. After all the cards in the machine have been listed and added, the machine automatically stops. An operator can handle two or more machines, as it is necessary merely to keep them supplied with cards. Cards pass through the machines at the rate of 2,500 per hour when there are large groups and about 2,000 per hour when there are small groups necessitating frequent summation.

"A study of the foregoing example of listing will indicate that the Tabulator-Printer is of great value in detecting errors. In the example given, the insured were all men of twenty years of age who had lapsed their policies with a duration of one year and were residing in State No. 16 (Illinois). Any errors resulting from putting cards wrongly in any of these categories would be easily detected. The same would hold true in other sortings, such as those among women who had died in the third insurance year. In that case column (c) should be '3'; column (d) should be 'F'; and column (e) should be 'D.' Again, if the group consisted of men aged 25 whose policies were existing at the close of the investigation and who had taken their policies in 1905 on the twenty payment life plan when they were resident in Massachusetts, column (a) should be '5'; column (b) should consist only of '25'; column (d) of 'M'; column (e) of 'E'; column (f) of '2'; column (g) of '25'; column (c) should be blank."

From the above descriptions of the mechanical systems in use it will be seen that their invention has made it possible to carry out expeditiously large mortality investigations which would involve extraordinary effort if carried out by hand. The successful operation of these systems, however, requires careful planning of every detail that is to be recorded.

In connection with these mechanical systems it should be kept in mind that in most cases it would be impracticable to punch the cards directly from the company's records. It is customary, therefore, to prepare preliminary sheets on which can be written the data necessary for the cards. The duration or age at exit may be calculated from the original records by the person making out the working sheet, thus reducing the punching of the cards to a simple and rapid operation. The form of sheet used in the American-Canadian Mortality Investigation is shown on the following page.

This form should be compared with the specimen card which is shown on page 27 and which was punched from it.

In investigations carried out by means of written cards, no preliminary working sheets would ordinarily be necessary, as the cards could be written directly from the original records. It should be noted how the card on page 25, used in the British Offices' Experience, is made up conveniently for copying from the original records and then for computing the durations and ages.

Before tabulating the cards, whether they be on the written or punched form, it is desirable to have some rough check on their accuracy even though they have already been checked individually with the source from which the data were drawn. This is particularly advisable if the records of several companies are being combined. In the American-Canadian Mortality Investigation it is stated:

"It seemed essential to have some check on the data so as to determine whether a section of the experience relating to existing, lapsed or deaths had been omitted. So far as concerns the principal companies which furnished their data on perforated cards, very careful checks were made by these companies, including calculation of expected deaths by a standard table and comparison with the actual death losses. For the other companies the bureau tabulated the existing, discontinued and dead, according to age, and ascertained if they were reasonable. The death losses were also subdivided by insurance years and comparison made between the groups by ages and insurance years. This would probably have brought out any material error.

"All cases were investigated which showed that the age of entry was over 70 as the number of such might be few and therefore errors in perforating might affect the results. Where the age at death was over 100 the company reporting the case was asked to verify the original record."

CHAPTER V. TABULATING THE DATA.

Before sorting the cards on which the details have been recorded, careful plans should be made for obtaining the exposed to risk in the easiest manner possible. A continuous process is used by which E_{x+1} , the exposed to risk at age $x + 1$, is derived from E_x . The formula to be employed will depend upon the method of analyzing the exposures and deaths—policy, calendar, or life year method—and upon the character of the mortality table to be constructed.*

A. Policy Year Method.

Let us first consider the construction of an aggregate table by the policy year method.

Let n_x = New entrants at age x .

e_x = Existing at age x at close of observations.

σ_x = Those under observation at age x when the observations began.

w_x = Withdrawals at completed age x .

g_x = Aggregate of fractional durations beyond completed age x arising from the w_x withdrawals.

θ_x = Deaths occurring between ages x and $x + 1$.

Let it be assumed that w_x includes both the withdrawals and the matured. If it is desirable to have them separated for other purposes, separate symbols may be assigned to each, but the formulae will hold with the corresponding substitutions. It will be noticed that the completed age at death is used, *i.e.*, the age at issue plus the curtate duration. Therefore the θ_x deaths must be treated as exposed to risk to age $x + 1$.

Then

$$E_x = \sum_0^x n_x + \sum_0^x \sigma_x - \sum_0^x w_x + g_x - \sum_0^x e_x - \sum_0^{x-1} \theta_x$$

and

$$E_{x+1} = E_x + n_{x+1} + \sigma_{x+1} - w_{x+1} + (g_{x+1} - g_x) - e_{x+1} - \theta_x. \quad (1)$$

* The authors wish to acknowledge their indebtedness to Mr. Robert Henderson for his concise explanation of formulae and methods of treating ages and durations contained in his monograph "Mortality Laws and Statistics" published by John Wiley & Sons, Inc.

It will be seen that the use of g_x assumes that the exact duration method has been used. If nearest or mean durations are employed the withdrawals take place at integral ages and g_x becomes zero for all values of x .

In order to carry out this formula, the information may be recorded on one sheet similar to the following.

Age.	(1) New En- trants.	(2) Enter- ing Observation.	(3) = (1) + (2).	(4) With- draw- als.	(5) Exist- ing.	(6) Dead.	(7) = (4) + (5) + (6) Pre- vious line.	(8) $E_x = (3)$ $- (7) + (8)$ Previous line.	(9) $q_x = \frac{(6)}{(8)}$
15									
16									
17									
18									
19									

The cards would be sorted by the mode of entrance corresponding to columns (1) and (2), if they were not already so separated because of their different years of issue. Each set of cards would then be sorted by age at entrance into the observations and the totals for each age recorded. The next step would be to sort them all in accordance with the mode of exit and then by age at exit. In tabulating the number of withdrawals at age x we shall have the sum of the fractional exposures, g_x , to tabulate also if the exact duration method was used. If, however, we tabulate the expression $(w_x + g_{x-1} - g_x)$ in column (4), the value of column (7) will be $(w_x + g_{x-1} - g_x + e_x + \theta_{x-1})$, which is the quantity that enters negatively in equation (1).

This method of tabulating might be described as assuming that, of the w_x withdrawals at completed age x , we consider g_x as withdrawn at age $x + 1$ and $w_x - g_x$ at age x . A separate sheet or column might be provided for recording the values of g_x and of $g_{x-1} - g_x$. It will be noticed that this latter value may be negative, in which cases its absolute value is deducted from w_x .

Where select tables are to be constructed the exposed to risk must be tabulated for each age at issue separately and the formulae become

$$E_{[x]+t} = n_{[x]} + \sum_1^t \sigma_{[x]+t} - \sum_0^t w_{[x]+t} + g_{[x]+t} - \sum_1^t e_{[x]+t} - \sum_0^{t-1} \theta_{[x]+t}$$

and

$$E_{[x]+t+1} = E_{[x]+t} + \sigma_{[x]+t+1} - w_{[x]+t+1} + (g_{[x]+t+1} - g_{[x]+t}) - e_{[x]+t+1} - \theta_{[x]+t} \quad (2)$$

where exact durations have been used.

If the nearest or mean durations are taken, $g_{[x]+t}$ vanishes and we have

$$E_{[x]+t+1} = E_{[x]+t} + \sigma_{[x]+t+1} - (w_{[x]+t+1} + e_{[x]+t+1} + \theta_{[x]+t}) \quad (3)$$

In all of the foregoing equations it has been assumed that observations began and ended on policy anniversaries in the first and last calendar years of the investigation, so that there are no fractional durations for σ_x and e_x ; otherwise the fractional exposures must be treated in a manner similar to those arising from the withdrawals.

In order to produce the values of $E_{[x]+t+1}$ by equation (3) it will be seen that the sorted cards may be tabulated on sheets drawn up somewhat as follows, using a separate sheet for each age at issue.

Age at Issue					Number of Entrants	
Duration.	(1) Entering Observations.	(2) Withdrawals.	(3) Existing.	(4) Dead.	(5) Decrement (2) + (3) - (1) + Previous Line (4).	(6) Exposed to Risk (6) Previous Line - (5).
0						
1						
2						
3						
4						

The exposed to risk at duration 0 would be found by subtracting the withdrawals at that duration from the number of entrants shown at the top of the sheet. There would naturally be none entering the observations or existing at duration 0.

To have the cards in the proper order, they must be sorted by age at issue, then by mode of exit and finally by duration at exit. The cards for $\sigma_{[x]+t}$ must also be sorted by duration at entrance into the experience.

If the exact duration method is used for the withdrawals the cards should show and be sorted by the curtate duration. For each duration there must then be calculated the value of $g_{[x]+t}$. This number may be tabulated as withdrawing at duration $t + 1$ and the balance of the withdrawals, $(w_{[x]+t} - g_{[x]+t})$, may be placed opposite the curtate duration, t . This will permit $E_{[x]+t}$ to be computed in the same way as where mean or nearest durations were employed. In writing down the withdrawals on the sheet two entries will be made for each duration, except duration 0, because of the withdrawals thrown forward and those thrown back. It would therefore be well to have the column for withdrawals subdivided for each duration by a light horizontal line.

Having thus obtained the exposed to risk the select mortality rates may be computed, for

$$q_{[x]+t} = \frac{\theta_{[x]+t}}{E_{[x]+t}}.$$

The period of selection may also be determined by combining the data on the sheets for the several ages at entry as described in Chapter I. Where an aggregate table is wanted in addition to select tables, the deaths and exposed to risk for each attained age may be footed from the sheets and the ratios of these totals calculated.

It may be desired to obtain an ultimate or a truncated table only. If we let t be the maximum duration excluded from such an experience then every policy in force not more than t years would be eliminated and every other policy would be recorded as entering at age $x + t$. In other respects the same method could be followed and the same formulae used as for an aggregate table. To construct a true ultimate table, however, it would not be possible to determine t beforehand and it would be necessary to list the select data for that purpose. The method just mentioned is therefore practicable only for a truncated table where the experience to be excluded is fixed before the investigation is tabulated.

In the British Offices' Assurance Experience the withdrawals, that is lapses and surrenders, were designated by $w_{[x]+t}$, while the matured, called "Terminations," were treated separately under the symbol $T_{[x]+t}$. For both classes the nearest duration method was used, properly modified to include the days of grace

in the exposures. The value of $w_{[x]+t}$ was the sum of $W^{(b)}_{[x]+t-1}$ and $W^{(a)}_{[x]+t}$, namely the withdrawals in the t th and $(t+1)$ th policy years that were taken as withdrawing at integral duration t . Similarly $T_{[x]+t}$ represents the sum of $T^{(b)}_{[x]+t-1}$ and $T^{(a)}_{[x]+t}$. The symbol $G_{[x]+t}$ was used to represent the net movement, *i.e.*, those entering the exposures at duration t less those emerging.

Therefore $G_{[x]+t} = \sigma_{[x]+t} - (\theta_{[x]+t-1} + w_{[x]+t} + T_{[x]+t} + e_{[x]+t})$ and, as $\sigma_{[x]}$ was used to represent $n_{[x]}$ so that $G_{[x]} = \sigma_{[x]} - w_{[x]}$,

$$E_{[x]+t} = E_{[x]+t-1} + G_{[x]+t} \text{ or } E_{[x]+t} = \sum_{\tau=0}^{\tau=t} G_{[x]+\tau}.$$

A specimen working sheet is shown on page 37.

This table contains the data only for "New" assurances, namely, issues from 1863 on, and therefore $\sigma_{[x]+t}$ has a value only at duration 0, representing new issues.

In combining the select data for the construction of an aggregate table it was necessary to sum the several values for the same attained age. Thus $w_x = w_{[0]+x} + w_{[1]+x-1} + w_{[2]+x-2} + \dots + w_{[x-1]+1} + w_{[x]+0}$ and similarly for θ_x , σ_x , T_x and e_x . Then $G_x = \sigma_x - (\theta_{x-1} + w_x + T_x + e_x)$.

The corresponding working sheet is given on page 38.

For the construction of truncated aggregate tables the following equation was used,

$$E_x^{(t)} = E_x - (E_{[x]} + E_{[x-1]+1} + \dots + E_{[x-t+1]+t-1}),$$

where $E_x^{(t)}$ indicates the exposures at attained age x excluding the first t years of duration.

The American-Canadian Mortality Investigation furnishes an interesting example of a working sheet combining policies issued both before and during the period of investigation. The experience was based on amounts insured to the nearest hundred dollars. Mean durations were used (see page 39).

The years of issue shown in column (2) apply to the Existing in column (3). Similarly column (7) refers to the Entered in column (8). These years were inserted to make the tabulation as simple as possible. Column (10) was obtained by summing column (9) from the bottom up instead of working from the top down in the usual manner by subtraction. The text of this investigation gives the formula

$$E_{[x]+n} = E_{[x]+n-1} - (e + w + d - s)_{[x]+n},$$

"New" Assurances.							With Profits.			
Duration.	Entrants.	Deaths.	WITHDRAWALS.		TERMINATIONS.		Existing in 1893.	Total Decrement.	Net Movement.	EXPOSED TO Risk.
			Fractional Exposure.	Balance.	Fractional Exposure.	Balance.				
	At Exact Duration t .	$t - 1$.	$t - 1$.	t .	$t - 1$.	t .	At Exact Duration t .	(3) + (4) + (5) + (6) + (7) + (8).	(2) - (9).	$\sum_0^t (10)$.
Modifications at Exact Duration t .										
(t).	$\sigma_{[x]+t}$.	$\theta_{[x]+t-1}$.	$W^{(b)}_{[x]+t-1}$.	$W^{(a)}_{[x]+t}$.	$T^{(b)}_{[x]+t-1}$.	$T^{(a)}_{[x]+t}$.	$\epsilon_{[x]+t}$.	$\theta + w + T + \epsilon$.	$G_{[x]+t}$.	$E_{[x]+t}$.
(1).	(2).	(3).	(4).	(5).	(6).	(7).	(8).	(9).	(10).	(11).
(0)	7,141	125	125	7,016	7,016
1	16	147	431	808	1,402	1,402	5,614
2	17	123	190	620	950	950	4,664
3	20	74	104	587	785	785	3,879
4	25	38	68	469	600	600	3,279
5	13	22	56	420	511	511	2,768
6	15	30	34	356	435	435	2,333
7	20	19	31	262	332	332	2,001
8	14	19	22	277	332	332	1,669
9	5	11	17	210	243	243	1,426
10	7	5	10	170	192	192	1,234

ENDOWMENT ASSURANCE EXPERIENCE, 1863-1893
Aggregate Tables. Male Lives.
 ABSTRACT OF DATA, with Computation of Numbers, Exposed to Risk.

CLASS E.

AGE 40 = $[x] + t$.

SECTION NP.

"New" Assurances.					With Profits.					
Dura- tion.	Entrants.	Deaths.	WITHDRAWALS.		TERMINATIONS.		Existing in 1893.	Total Decrement.	Net Movement.	Exposed TO Risk.
			Fractional Exposure.	Balance.	Fractional Exposure.	Balance.				
	Curtate Duration.									
At Exact Duration t .	$t-1$	$t-1$	t	$t-1$	t	At Exact Duration t .	$(3)+(4)$ $+(5)+(6)$ $+(7)+(8)$.	$(2)-(9)$.	$\sum_0^t (10)$.	
	Modifications at Exact Duration t .									
(0).	$\sigma_{[x]+t}$.	$\theta_{[x]+t-1}$.	$W^{(b)}_{[x]+t-1}$.	$W^{(a)}_{[x]+t}$.	$T^{(b)}_{[x]+t-1}$.	$T^{(a)}_{[x]+t}$.	$\theta + w + T$ $+ \epsilon$.	$G_{[x]+t}$.	$E_{[x]+t}$.	
(1).	(2).	(3).	(4).	(5).	(6).	(7).	(9).	$+$ (10).	(11).	
(0)	1,915	17	17	1,898	...	1,898
1	2	4	22	77	320	318	1,557
2	2	11	27	38	289	287	1,524
3	2	8	14	29	260	258	1,341
4	1	8	11	23	234	233	1,327
5	1	9	10	26	...	2	269	268	1,333
6	1	9	10	20	249	248	1,250
7	1	14	9	12	214	213	1,278
8	1	11	5	12	199	198	1,113
9	1	9	7	11	...	1	175	174	1,025
10	9	9	11	...	2	167	167	944

AMERICAN-CANADIAN MORTALITY INVESTIGATION.
1900-1915.
American Men.

FROM THE RECORDS OF INSURED LIVES.

39

The Unit of Amount is \$100. AGE AT ENTRY 40.										
(1). Insurance Year.	(2). Issues of.	(3). Ex- isting.	(4). With- drawn.	(5). Dead.	(6). Cols. (3) +(4) + (5).	(7). Issues of.	(8). Entered.	(9). Col. (6) Minus Col. (8).	(10). Summation of Col. (9) Exposed to Risk.	(11). Col. (5) ÷ Col. (10) Rate of Mortality.
1	1914	\$5,902.3	\$15,996.6	\$258.3	\$22,157.2	1899	\$7,442.4	\$14,714.8	\$67,143.9	.00385
2	1913	5,049.4	3,908.4	210.2	9,168.0	1898	6,244.1	2,923.9	52,429.1	.00401
3	1912	4,647.1	2,219.8	253.5	7,120.4	1897	4,910.7	2,209.7	49,505.2	.00512
4	1911	3,429.8	2,031.5	274.0	5,735.3	1896	3,754.3	1,981.0	47,295.5	.00579
5	1910	3,199.5	1,669.4	223.2	5,092.1	1895	3,284.4	1,807.7	45,314.5	.00493
6	1909	3,045.8	1,412.6	329.7	4,788.1	1894	2,636.7	2,151.4	43,506.8	.00758
7	1908	1,938.5	1,248.2	371.2	3,557.9	1893	1,987.7	1,570.2	41,355.4	.00898
8	1907	1,379.2	902.6	317.6	2,599.4	1892	2,293.7	305.7	39,785.2	.00798
9	1906	1,858.6	976.2	318.3	3,153.1	1891	2,042.2	1,110.9	39,479.5	.00806
10	1905	1,768.6	2,490.8	407.2	4,666.6	1890	1,966.8	2,699.8	38,368.6	.01061
—	—	—	—	—	—	—	—	—	—	—
51	1864	6.0	29.0	35.0	1849	3.5	31.5	71.1	.40788
52	1863	6.0	1.0	12.6	19.6	1848	2.0	17.6	39.6	.31818
53	1862	1.0	4.0	5.0	1847	5.0	22.0	.18182
54	1861	5.0	5.0	1846	5.0	22.0	.22727
55	1860	3.5	3.5	1845	3.5	17.0
56	1859	7.0	7.0	1844	7.0	13.5	.51852
57	1858	1.5	1.5	1843	1.5	6.5
58	1857	5.0	5.0	1842	5.0	5.0
59	1856	1841
60	1855	1840
Totals		\$61,183.8	\$45,538.8	\$16,344.1	\$123,066.7		\$55,922.8	\$67,143.9	\$884,778.8	

but the arithmetical work followed the equation

$$E_{[x]+n-1} = E_{[x]+n} + (e + w + d - s)_{[x]+n}.$$

In these equations $d_{[x]+n}$ refers to the deaths in the n th insurance year, and not to the deaths at curtate duration n .

Mr. W. J. H. Whittall (*J. I. A.*, Volume XXXI, 167) introduced a convenient notation to show in any formula for the exposed to risk what methods were used in determining the age at entry and the duration. Thus

$$\begin{aligned} x &= \text{Exact age} \\ \{x\} &= \text{Age next birthday} \\ (x) &= \text{Nearest age} \\ |x| &= \text{Mean age} \\ t &= \text{Exact duration} \\ \overline{t} &= \text{Curtate duration} \\ (\overline{t}) &= \text{Nearest duration} \\ |t| &= \text{Mean duration} \end{aligned}$$

To the age symbols might be added \underline{x} to indicate the completed age, or age last birthday, so that $\{x\} = \underline{x} + 1$.

This notation is appropriate for aggregate exposures, and for calendar year and life year as well as policy year investigations, but a difficulty arises in applying the age symbols to select notation where the age at entry is indicated by $[x]$. This was recognized by the author who stated

"It will be understood that in dealing with, for instance, age at exit $\{x\} + |t|$, each symbol is variable, and the expression includes entrants of different ages and assured of different durations, subject only to the total ages attained being correct according to the terms of the formula."

Even in select notation, however, the duration symbols may be used.

B. Calendar Year Method.

The calendar year method, as previously stated, is not directly applicable to the construction of select tables. In considering the formula for the exposed to risk in an aggregate table, it is essential to define the symbols carefully.

Let n_x = New entrants aged x on January 1st of the year of entry.

e_x = Existing at age x on December 31st of the last year covered by the investigation.

w_x = Withdrawals in calendar year beginning at age x .

θ_x = Deaths in calendar year beginning at age x .

σ_x = Entering observations at age x on January 1st of the first year covered by the investigation.

Let it further be assumed that an investigation of the data justifies the use of the fraction f as the part of a calendar year elapsed at the average date of entry, and the fraction g as the average part of a calendar year elapsed at the date of withdrawal.

Then

$$E_x = \sum_0^x n_x - f n_x + \sum_0^x \sigma_x - \sum_0^x w_x + g w_x - \sum_0^x e_x - \sum_0^{x-1} \theta_x,$$

$$E_{x+1} = E_x + n_{x+1} - f(n_{x+1} - n_x) + \sigma_{x+1} - w_{x+1} + g(w_{x+1} - w_x) - e_{x+1} - \theta_x. \quad (4)$$

It is usual to assume that $f = g = \frac{1}{2}$, in which case this equation is reduced to

$$E_{x+1} = E_x + \frac{n_{x+1} + n_x}{2} + \sigma_{x+1} - \frac{w_{x+1} + w_x}{2} - e_{x+1} - \theta_x. \quad (5)$$

If a truncated table were to be constructed, eliminating the experience of an integral number of years, t , from issue, the same equations would apply; but the age at entry would be $x + t$ instead of x and all policies terminated for any cause at the end of t years or before would be eliminated from the experience.

When the age at entry is taken as the nearest age or mean age at issue, the exposures and deaths are given for half ages instead of integral ages as mentioned in Chapter III. Thus n_x could be described as the new entrants aged $x + \frac{1}{2}$ at date of issue, $x + \frac{1}{2}$ being integral. Equations (4) and (5) are still applicable, but the ages for all the symbols must be taken as half ages.

In forming the H^M table the exposed to risk were calculated in the following manner:

$$E_x = E_{x-1} + \frac{1}{2}(n_x + n_{x+1}) - \frac{1}{2}(w_x + w_{x+1}) - (d_x + e_x).$$

Where* n_x = Entrants at age x next birthday.

* As the experience covered all policies from their dates of issue, there were no cases entering observation after issue.

$$\begin{array}{l}
 w_x = \text{Withdrawals at age } x \text{ next birthday} \\
 d_x = \text{Deaths at age } x \text{ next birthday} \\
 e_x = \text{Existing at age } x
 \end{array}
 \left\{ \begin{array}{l}
 \text{i.e., age at} \\
 \text{entry next} \\
 \text{birthday plus} \\
 \text{mean dura-} \\
 \text{tion = age} \\
 \text{at end of cal-} \\
 \text{endar year of} \\
 \text{exit}
 \end{array} \right\}.$$

This formula assumes that the observations close at the end of a calendar year.

In investigating the data of the Scottish offices which contributed to the Institute Experience, Mr. Meikle transformed this formula as follows:

$$\begin{aligned}
 E_x &= \sum_0^x n_x - \sum_0^x (d_x + w_x + e_x) + \frac{1}{2}(n_{x+1} - w_{x+1}) \\
 &= \sum_0^x n_x - \sum_0^x f_x + \frac{1}{2}(n_{x+1} - w_{x+1})
 \end{aligned}$$

where $f_x = d_x + w_x + e_x$. Mr. Ryan (*J. I. A.*, XXVI, 257) points out that Mr. Meikle's form is easier to apply than the Institute form, but suggests a further modification to the following:

$$E_x = \sum_0^x (n_x - f_x) + \frac{1}{2}(n_{x+1} - w_{x+1}).$$

This would save one step in the calculations.

It should also be kept in mind that the data entering into this experience were separated by age at entry, in order that the effects of selection might be seen and approximate select tables prepared at a later date. The following illustration will show the arrangement of the figures:

* Current Age at Entry 30.

Number of Entrants 5791.

Current Age at Exit.	Existing.	Discontinued.	Died.
30	319	75	4
31	252	365	28
32	230	220	35
33	235	153	49

*Current Age means age next birthday.

The calendar year of entry, called "Year 0," gives the exposed to risk for age 29. The next calendar year, or "Year 1" refers to age 30 and so on. Thus by the Institute formula for this one age at entry

$$E_{29} = \frac{5791}{2} - \frac{75}{2} = 2858,$$

$$E_{30} = E_{29} + \frac{5791}{2} - \frac{75 + 365}{2} - (4 + 319) = 5210.5.$$

In forming the H^m table the figures were combined for each attained age, irrespective of age at entry.

C. Life Year Method.

As the life year method is not suitable for the construction of select tables, consideration will be given only to aggregate tables. Let us suppose that exact ages at entrance and withdrawal have been tabulated.

Then let n_x = New entrants at age x last birthday.

w_x = Withdrawals at age x last birthday.

θ_x = Deaths at age x last birthday.

Let it also be assumed that the aggregate of the fractions of a year since last birthday at the time of entry of n_x are tabulated and denoted by f_x , and that g_x is the sum of the fractions of exposures in the year of withdrawal among w_x .

Then

$$E_x = \sum_0^x n_x - f_x - \sum_0^x w_x + g_x - \sum_0^{x-1} \theta_x \quad (6)$$

$$E_{x+1} = E_x + n_{x+1} - (f_{x+1} - f_x) - w_{x+1} + (g_{x+1} - g_x) - \theta_x.$$

If instead of using f_x and g_x , which are the sums of the fractions and vary with the attained age, average fractions per entrant and withdrawal, f and g , are adopted for all ages, we have

$$E_x = \sum_0^x n_x - fn_x - \sum_0^x w_x + gw_x - \sum_0^{x-1} \theta_x,$$

$$\begin{aligned} E_{x+1} &= E_x + n_{x+1} - f(n_{x+1} - n_x) - w_{x+1} \\ &\quad + g(w_{x+1} - w_x) - \theta_x \\ &= E_x + \{fn_x + (1 - f)n_{x+1}\} \\ &\quad - \{gw_x + (1 - g)w_{x+1}\} - \theta_x. \end{aligned} \quad (7)$$

If there are persons under observation when the period of exposure began and existing at the end of that period, the former

may be treated as new entrants at their ages at which they came into the experience and the latter in the same manner as withdrawals. In such a case they must of course be taken into account in arriving at the average fractions f and g .

It has been stated in Chapter III that the ages at entry and withdrawal may be taken as the mean ages. It might be contended that we then have a calendar year method. This is not the case, however, as under the latter method the deaths are analyzed by calendar years, *i.e.*, by age at the beginning of the calendar year of death. Under the life year method the deaths are analyzed by the age last birthday at death, and mean ages at entry and withdrawal are used simply as a device for getting rid of fractional exposures at such times.*

If the observations begin and close with persons under observation or existing, their ages at such times could not ordinarily be taken as mean ages. The age nearest birthday might be used. When, however, such exposures commence and cease at the end of two different calendar years, another convenient method may be adopted. Let σ_x represent those entering on January 1st at the start, at mean age x in the preceding calendar year. They will be exposed for half a year at age x as we may assume that on the average they are aged $x + \frac{1}{2}$ on January 1st, when the exposures begin. Similarly e_x represents the existing on the final December 31st, age x being the mean age in the year just ended. There will therefore be on the average a half year of exposure at age x . Then if x represents the mean age in the symbols n_x and w_x , while θ_x represents the deaths at age x last birthday, we have

$$E_x = \sum_0^x n_x + \sum_0^x \sigma_x - \frac{1}{2}\sigma_x - \sum_0^{x-1} \theta_x - \sum_0^x w_x - \sum_0^x e_x + \frac{1}{2}e_x,$$

$$E_{x+1} = E_x + n_{x+1} + \frac{1}{2}(\sigma_x + \sigma_{x+1}) - \theta_x - w_{x+1} - \frac{1}{2}(e_x + e_{x+1}). \quad (8)$$

If the period of investigation had not extended over exact calendar years, equation (8) could still be used if x in the symbols σ_x and e_x were the age last birthday, for we should still have approximately one half year of exposure at age x in both cases.

* Mr. W. J. H. Whittall (*J. I. A.*, Vol. XXXI, 163) discusses the definition of methods and states (p. 164): "It would, therefore, seem probable that in the system of determining the ages at death in any particular formula will be found the key to its proper classification."

The last method of estimating exposures mentioned in Chapter III is to take the ages at entry and exit as the ages nearest birthday, except for the age at death. This produces no fractional exposures, as those under observation at the start or existing at the close of the period may be likewise treated and included in the new entrants and withdrawals respectively.

Then

$$E_x = \sum_0^x n_x - \sum_0^{x-1} \theta_x - \sum_0^x w_x, \quad (9)$$

$$E_{x+1} = E_x + n_{x+1} - \theta_x - w_{x+1}.$$

One example of the life year method is the experience of the Economic Office, 1862, taken out by J. J. Downes, who was the introducer of the card system. A description of the methods employed is given by W. J. H. Whittall, *J. I. A.*, Vol. XXXI, 171. Exact ages to four decimal places were computed at entrance and exit, which is an unnecessary refinement. The method of tabulation was also laborious.

Mr. A. J. Finlaison used the life year method in compiling the Government Annuitants' Experience, 1883. After a test he assumed that on the average, four months had elapsed at date of entry since the last birthday. Accordingly each new entrant was assumed to be exposed to risk for two-thirds of a year at the age last birthday at entry. Ages at exit were taken accurately, the exposures on the existing terminating on exact birthdays. Mr. Finlaison used the data to form select tables of annuitants' mortality by assuming that the eight months mortality rate following entry was the rate for the first year after entry, and so on. Thus the approximate periods of experience $0-\frac{2}{3}$, $\frac{2}{3}-1\frac{2}{3}$, $1\frac{2}{3}-2\frac{2}{3}$, etc., were considered to be annuity years 1, 2, 3, etc. This illustrates how, in attempting to get mortality rates for each exact year of life, the advantages of accurate tables by policy years must be sacrificed. As the effect of self-selection by annuitants is often marked for the first few policy years, the understatement of the mortality in the early policy years under Mr. Finlaison's method would appear to be more important than the loss of accuracy that would be introduced by the use of the policy year method. The annuity experience of the British Offices, 1863-1893, was based on annuity years, the age at entry being the age nearest birthday at date of purchase.

CHAPTER VI. TABLES WITH MORE THAN ONE DECREMENTAL FACTOR.

A mortality table, as heretofore considered, is a table showing the number living at each age, either in the select or in the aggregate form. This number decreases as the age increases only because of the deaths. Death is the only decremental factor in such a table. We may, however, have a table which has one or more additional decremental factors, such as those of withdrawal, marriage or the re-marriage of widows. If the additional factor is that of withdrawal, then the table will show for each age the number living who have not withdrawn, and in proceeding from one age to the next there is deducted the number who have withdrawn and those who have died without withdrawing during that year of age. There is given below a specimen table (*J. I. A.*, Vol. XXXIII, 196) in the select form for age 20 at entry:

Mortality Table, Showing Mortality and Withdrawal.				
Dura- tion.	Numbers Living.	Deaths.	With- drawals.	Deaths and With- drawals.
(<i>t</i>).	$l_{[x]+t}$.	$d_{[x]+t}$.	$w_{[x]+t}$.	$(d + w)_{[x]+t}$.
0	100,000	426	14,970	15,396
1	84,604	373	10,549	10,922
2	73,682	335	7,717	8,052
3	65,630	307	5,892	6,199
4	59,431	297	4,741	5,038

Before starting to analyze the experience from which it is planned to construct such a table we must examine the functions by which we may compute the successive figures in the table. The rate of withdrawal will ordinarily vary with the duration of insurance as well as with age and therefore select tables are best adapted to display the results. It may also be noted that the rate usually varies with the plan of insurance, making separate tables by plan desirable in many instances. The investigation might be based on lives, policies, or amounts insured.

Let us adopt the following notation:

Table Notation:

$l'_{[x]+t}$ = the survivors at age $x + t$ of the entrants at age x who have not withdrawn.

$d'_{[x]+t}$ = the deaths during the $(t + 1)$ th year, before withdrawal.

$w'_{[x]+t}$ = the withdrawals during the $(t + 1)$ th year.

Investigation Notation:

$E_{[x]+t}$ = the exposed to risk of death.

$(wE)_{[x]+t}$ = the exposed to risk of withdrawal, the deaths being given their respective fractional durations while the withdrawals are given a full year's exposure.

$\theta_{[x]+t}$ = the deaths in the $(t + 1)$ th year.

$w_{[x]+t}$ = the withdrawals in the $(t + 1)$ th year.

It may be noticed that the definition of $(wE)_{[x]+t}$ is consistent with the method by which $E_{[x]+t}$ is determined, as in the latter case the withdrawals are given their proper fractional exposures but the deaths are counted as exposed to risk for the entire year of death. Thus to find the proportion that withdraw of those persons who are exposed to withdrawal for a year, we must count each person who withdraws as a unit and not a fraction in the exposures.

In tabulating $w_{[x]+t}$ the question arises whether the withdrawals at the exact end of a policy year, which are generally a large proportion of the total withdrawals, should be grouped with those during the year just closed or with those of the subsequent year. If we consider the first policy year it will be seen that the former course must be adopted in order to get a full year of withdrawals recorded in that year. We shall then have the rates of withdrawal during and at the end of each year, and $l'_{[x]+t}$ will represent the survivors of the original entrants who continue their policies beyond t years from issue. It should be remembered, however, that not all the $l'_{[x]+t}$ persons continue their policies to the end of the $(t + 1)$ th year, for some will withdraw at fractional durations.

If we now consider the double decrement table we see that the

probability that a life aged $[x] + t$ will withdraw within a year is

$$(wq)_{[x]+t} = \frac{w'_{[x]+t}}{l'_{[x]+t}},$$

which is different from the rate of withdrawal, among those who are alive. This latter rate is

$$(wr)_{[x]+t} = \frac{w'_{[x]+t}}{l'_{[x]+t} - \frac{1}{2}d'_{[x]+t}}$$

on the assumption of a uniform distribution* of deaths, and is the rate given in an investigation by the ratio $w_{[x]+t}/(wE)_{[x]+t}$. Similarly the probability of a life dying before withdrawal becomes

$$q'_{[x]+t} = \frac{d'_{[x]+t}}{l'_{[x]+t}},$$

whereas the rate of mortality, on the assumption of a uniform distribution of withdrawals, is

$$q_{[x]+t} = \frac{d'_{[x]+t}}{l'_{[x]+t} - \frac{1}{2}w'_{[x]+t}}.$$

The value of $q_{[x]+t}$ is given in the investigation by $\theta_{[x]+t}/E_{[x]+t}$. It is evident that in a single decrement mortality table q' becomes q , because the column of survivors, l , is reduced only by deaths.

In order to construct a table from the data we must find relations by which, starting with a given radix, we may find the values of $d'_{[x]+t}$ and $w'_{[x]+t}$. We can then construct the l' column by the equation $l'_{[x]+t+1} = l'_{[x]+t} - d'_{[x]+t} - w'_{[x]+t}$.

The simplest procedure would appear to be to calculate the values of q' and (wq) directly from the data, for these probabilities when multiplied by l' give us d' and w' . If this were done we should find the probabilities from the data as follows.

$$q'_{[x]+t} = \frac{\theta_{[x]+t}}{E'_{[x]+t}}, \quad (10)$$

$$(wq)_{[x]+t} = \frac{w_{[x]+t}}{(wE')_{[x]+t}}, \quad (11)$$

where E' and (wE') are obtained by giving $\theta_{[x]+t}$ and $w_{[x]+t}$ a

* It should be noticed that this assumption as applied to a double decrement table is not equivalent to the assumption of a uniform distribution of deaths in a single decrement table.

full year's exposure in both cases, and therefore

$$E'_{[x]+t} = (wE')_{[x]+t}.$$

Using the notation for single decrement tables except for the change in the definition of $w_{[x]+t}$, and assuming that the period of observation begins and ends on policy anniversaries, we have

$$E'_{[x]+t} = n_{[x]} + \sum_0^t \sigma_{[x]+t} - \sum_0^{t-1} w_{[x]+t} - \sum_0^t T_{[x]+t} - \sum_0^{t-1} \theta_{[x]+t} - \sum_1^t e_{[x]+t}$$

where $T_{[x]+t}$ represents the matured, such as endowments and terms, taken to mean or nearest durations. Then

$$E'_{[x]+t+1} = E'_{[x]+t} + \sigma_{[x]+t+1} - (w_{[x]+t} + \theta_{[x]+t} + T_{[x]+t+1} + e_{[x]+t+1}). \quad (12)$$

This gives us the denominator common to equations (10) and (11).

It frequently happens that the probabilities q' and (wq) are not as useful as the rates q and (wr) . If we wish to compare the results of the investigation with those of other experiences the latter rates should be used as q is independent of (wr) and vice versa, while q' and (wq) are each functions of both the mortality and the withdrawal rates. It may be desired to combine the withdrawal rate from one experience with the mortality rate of another, in which event we need to have the independent rates, q and (wr) .

Let us then see how we can construct the double decrement table from these rates.

On the assumption of a uniform distribution of deaths and withdrawals during policy years we have

$$\begin{aligned} (l' - \tfrac{1}{2}w')q &= d', \\ (l' - \tfrac{1}{2}d')(wr) &= w'. \end{aligned}$$

Solving the equations we have

$$d' = \frac{l'q[1 - \tfrac{1}{2}(wr)]}{1 - \tfrac{1}{4}q(wr)}, \quad (13)$$

$$w' = \frac{l'(wr)[1 - \tfrac{1}{2}q]}{1 - \tfrac{1}{4}q(wr)}. \quad (14)$$

We may thus compute the decrements from the rates q and (wr) , which can be obtained from the deaths, withdrawals, and exposed to risk as previously shown. We can also construct a table where q is derived from one experience and (wr) from

another. We may also obtain a check formula for the l' column, as follows:

$$\begin{aligned}
 l_{x+1}' &= l' - d' - w' \\
 &= l' \left[1 - \frac{q + (wr) - q(wr)}{1 - \frac{1}{4}q(wr)} \right] \\
 &= l' \frac{(1 - q)[1 - (wr)] - \frac{1}{4}q(wr)}{1 - \frac{1}{4}q(wr)}.
 \end{aligned} \tag{15}$$

These equations were developed by Dr. T. B. Sprague in a paper on marriage and mortality tables (*J. I. A.*, Vol. XXI, 406) and were applied to tables of mortality and withdrawal by Mr. T. G. Ackland (*J. I. A.*, Vol. XXXIII, 194). Mr. Ackland, in his paper, discusses very fully the use of exact, mean and nearest durations in obtaining the exposed to risk, $E_{[x]+t}$ and $(wE)_{[x]+t}$, and concludes that the nearest duration method is, upon the whole, the best suited for obtaining the rates of mortality and withdrawal from a large body of insured lives.

Mr. Ackland deduces the formulae for the exposed to risk where the period of investigation extends over integral policy years in the following manner.

Let $s_{[x]+t}$ = the survivors in force at the commencement of the period of observation, at integral duration t .

$n_{[x]}$ = the new entrants at age x .

$e_{[x]+t}$ = the cases existing at the close of the period of observation, at integral duration t .

$(aw)_{[x]+t}$ = the withdrawals having a duration greater than t that are treated as having integral duration t , *i.e.*, those with a duration less than $t + .5$ and one-half those having a duration of exactly $t + .5$.

$(bw)_{[x]+t}$ = the balance of the withdrawals with duration greater than t but not in excess of $t + 1$. These are treated as having integral duration $t + 1$.

Then $(aw)_{[x]+t} + (bw)_{[x]+t-1} = w_{[x]+t}$, the total withdrawals at integral duration t by the nearest duration method.

$(ad)_{[x]+t}$ = deaths in first half of $(t + 1)$ th year,

$(bd)_{[x]+t}$ = deaths in last half of $(t+1)$ th year, and
 letting $d_{[x]+t} = (ad)_{[x]+t} + (bd)_{[x]+t-1}$ *

also, let

$$(e + w + d)_{[x]+t} = F_{[x]+t}$$

$$(s - F)_{[x]+t} = G_{[x]+t}$$

where

$$t = 0, \quad G_{[x]} = n_{[x]} - F_{[x]}.$$

Then

$$E_{[x]+t} = \sum_0^t (G) + (ad)_{[x]+t} \quad (16)$$

$$(wE)_{[x]+t} = \sum_0^t (G) + (aw)_{[x]+t} \quad (17)$$

On page 52 is given a copy of the working sheet used to carry out the tabulation for age 20 at entry.

Mr. Ackland's notation is retained in this table. It must be carefully noted that $w_{[x]+t}$ in columns (4) and (13) are not the same functions, and a like condition is true of $d_{[x]+t}$ shown in columns (5) and (10). Column (11) is obtained by dividing (10) by (9), and column (14) is the quotient of $(13) \div (12)$. The brackets in column (5) indicate how the numbers in columns (4) and (5) were combined to obtain (6). Columns (10) and (13) were obtained from (5) and (4) respectively by combining the figures set down opposite each duration.

The data being analyzed had been obtained from a Clerks' Association in which subscriptions were payable monthly on the first day of each month. Therefore the withdrawals could be assumed to be distributed uniformly over each year of duration. Mr. Ackland found that the exact average fractional duration at withdrawal was approximately one half year. Accordingly formulae (13), (14) and (15) can be considered sufficiently accurate.

Dr. Sprague, in the paper previously mentioned, also developed for a double decrement table central rates corresponding to

$$m = \frac{d'}{l' - \frac{w'}{2} - \frac{d'}{2}},$$

* It will be seen that $d_{[x]+t}$ has not the usual significance of $\theta_{[x]+t}$ or $d_{[x]+t}$ but is calculated similarly to $w_{[x]+t}$ in accordance with nearest durations.

Dura- tion. viving.	Ex- isting.	With- drawals.	Deaths.	Total Decre- ment.	Net Move- ment.	$\bar{E}_{[x]+t}$ $= \sum_0^t (G)$	Mortality.			Withdrawal.			Dura- tion.
							Exposed.	D'ths.	Rate.	Exposed.	With- draw- als.	Rate.	
$t.$	$s_{[x]+t}.$	$(bw)_{[x]+t-1}$ $(aw)_{[x]+t}$ $= w_{[x]+t}.$	$(bd)_{[x]+t-1}$ $(ad)_{[x]+t}$ $= d_{[x]+t}.$	$(e+w+d)$ $= F_{[x]+t}.$	$(s-F)$ $= G_{[x]+t}.$	$\bar{E}_{[x]+t}$ $= \sum_0^t (G)$	$E_{[x]+t}$ $= \sum_0^t (G)$ $+ (ad).$	$d_{[x]+t}.$	$q_{[x]+t}$ $= \frac{d}{\bar{E}}.$	$(wE)_{[x]+t}$ $= \sum_0^t (G)$ $+ (aw).$	$w_{[x]+t}.$	$(wq)_{[x]+t}$ $= \frac{w}{(wE)}.$	$t.$
(1) 0	(2) 332	(4) 21 14	(5)	(6) 21	(7) +311	(8) 311	(9) 311	(10) ..	(11)	(12) 332	(13) 35	(14) .1054	(15) 0
1	128	85 23 17	0 3	122	+ 6	317	317	3	.00946	340	40	.1177	1
2	93	7 19 17	2 0	48	+ 45	362	364	2	.00549	381	36	.0945	2
3	124	75 24 16	1 0	117	+ 7	369	370	1	.00270	393	40	.1018	3
4	116	80 28 14	2 0	126	- 10	359	361	2	.00554	387	42	.1085	4
5	121	77 14 18	2 0	107	+ 14	373	375	2	.00533	387	32	.0827	5

* In terms of our notation the heading for this column should be (wr) instead of (wq) .

and

$$(wm) = \frac{w'}{l' - \frac{w'}{2} - \frac{d'}{2}},$$

which are mutually independent but from which can be computed, if desired, the rates q and (wr) by the relations

$$q = \frac{2m}{2 + m}$$

and

$$(wr) = \frac{2(wm)}{2 + (wm)}.$$

Central rates were employed by Mr. Ackland, who shows that

$$d' = l' \frac{m}{1 + \frac{m + (wm)}{2}}$$

and

$$w' = l' \frac{(wm)}{1 + \frac{m + (wm)}{2}}.$$

and also that

$$l_{+1}' = l' \frac{1 - \frac{m + (wm)}{2}}{1 + \frac{m + (wm)}{2}}.$$

In computing the central rates he divided $\theta_{[x]+t}$ and $w_{[x]+t}$ by the adjusted exposed to risk, $\bar{E}_{[x]+t}$, in each case. $\bar{E}_{[x]+t}$ was the number exposed to risk computed up to the actual cessation of the risk from any cause—under the nearest duration method, to the nearest integral year. This makes it unnecessary to calculate separate exposed to risk for deaths and withdrawals. These equations are, however, incorrect where there is not a uniform distribution of withdrawals.

A life insurance company which issues policies only with annual, semi-annual and quarterly premiums could not assume a uniform distribution of withdrawals. For example, assuming that lapse is as likely at the second semi-annual premium as at the first, the average fractional duration would be

$$\frac{1}{2} \left(\frac{1}{2} + 1 \right) = \frac{3}{4},$$

and similarly for quarterly lapses

$$\frac{1}{4} \left(\frac{1}{4} + \frac{1}{2} + \frac{3}{4} + 1 \right) = \frac{5}{8}.$$

If we assume $62\frac{1}{2}$ per cent. of the lapses are on policies issued at annual rates, $32\frac{1}{2}$ per cent. at semi-annual, and 5 per cent. at quarterly, the average fractional duration for all lapses would be

$$.625 + .325 \times \frac{3}{4} + .05 \times \frac{5}{8} = .9.$$

If we felt warranted in considering that this represented the proper average fraction for all withdrawals, including surrenders, at all durations, we might record $w_{[x]+t}$ in column (4) as equal to $.9w_{[x]+t-1} + .1w_{[x]+t}$, *i.e.*, refer nine-tenths of the withdrawals during or at the end of any year to the integral duration at the end of the year, and one-tenth to the beginning of the year, instead of tabulating by the nearest duration method the functions (*aw*) and (*bw*).

This example points out the error in assuming in such a case that

$$\left(l' - \frac{1}{2}w' \right) q = d'.$$

Continuing the above illustration it would be necessary to change this equation to

$$(l' - .1w')q = d',$$

which, together with

$$\left(l' - \frac{1}{2}d' \right) (wr) = w',$$

might be solved to give d' and w' in terms of l' , q and (wr) .

In the discussion of Mr. Ackland's paper, Mr. R. Todhunter presented a new view of withdrawal rates. He states:

"I suggest, in the first place, that discontinuances in ordinary business are attributable to a force essentially different in character from the force which operates to produce mortality. Death claims are caused by a force which operates continuously throughout the history of policies, whereas discontinuances are caused—so far as regards the large majority of policies—by a discontinuous force coming into operation at certain recurring epochs. A policyholder does not exercise his option of withdrawing continuously in the same sense that he is continuously subject to the risk of death; he exercises it periodically, in most cases on the occasions of his receiving a renewal notice,

and having to decide whether to pay or not to pay. This applies, I think, to nearly all lapses, the only exceptions that occur to me being the few cases in which a policyholder pays a series of fines to extend the days of grace for successive short periods; and lapses constitute a very large proportion of the whole number of discontinuances. It applies, also, to all surrenders carried out at or about the renewal date. The proportion which these form of the entire body of surrenders, will no doubt vary in different classes of business; in a small experience of two years, I found that 55 out of a total of 94 surrenders took place during the days of grace, 11 took place within a month before the renewal date (probably on receipt of renewal notices), and the remaining 28 were scattered. Having regard to the small proportion that the number of scattered surrenders forms of the entire number of lapses and surrenders, I think it may fairly be stated as a general proposition, that discontinuances are mainly due to the exercise of a periodical option at or about the renewal date.

"If this proposition be admitted, it follows that the force which causes discontinuances would be more appropriately measured by rates of non-renewal, than by rates or forces of withdrawal—in other words, by the ratios that the withdrawals at definite epochs bear to the exposed to risk of withdrawal at those epochs, than by the ratios that the withdrawals in given periods bear to certain numbers supposed to be continuously exposed to the risk of withdrawal throughout those periods. As applied to a collected experience, this second proposition presupposes a policy-year tabulation of the observed facts. A tabulation by calendar-years, or years of life, will, of course, have the effect of spreading the discontinuances over the years of observation, and will thus exhibit something of the nature of a continuous force of discontinuance. If it be admitted that such a force has no real existence, the fact that it is artificially created by any method of tabulation other than one that follows the years of assurance may be considered another argument in favour of the Policy-year Method."

The sort of table suggested can be clearly seen from that given below.*

Year of Assurance.	Exposed to Risk of Death.	<i>d.</i>	<i>q.</i>	Exposed to Risk of Non-Renewal.	<i>w.</i>	Rate of Non-Renewal.
1	8,016	30	.0037	7,986	1,288	.161
2	6,313	36	.0057	6,277	476	.076
3	5,535	25	.0045	5,510	297	.054
4	4,945	29	.0059	4,916	149	.030
5	4,404	25	.0057	4,379	176	.040
6	3,958	15	.0038	3,943	95	.025
Total	33,171	160	.0048	33,011	2,481	.075

The nearest duration method is a convenient one for arranging the withdrawal data in the proper form for determining the rate

* Terminations other than by death and withdrawal are not shown in this table.

of non-renewal. Each of the numerous withdrawals at the end of the policy year would be located exactly. Those at semi-anniversaries would be alternately located at the beginning and end of the year, lapses at the end of the first and third quarters would be thrown to the nearest anniversary as would also the usually small proportion of surrenders that are scattered through the year. The result would be that all withdrawals would be considered as taking place on policy anniversaries. This distorts the facts somewhat, but on the whole gives a more nearly correct view than to assume that all the withdrawals are spread throughout the policy year. Mr. Todhunter suggests that it may be desirable in a large experience to exhibit the rates of mortality and non-renewal by quarters of a year for the first two policy years.

It may also be seen that a double decrement table constructed in this manner is a useful one for determining the present value of future commissions, because $l'_{[x]+t}$ will be the number of survivors at duration t who pay premiums to the end of the $(t + 1)$ th year. The table could also be used for comparative purposes and adjusting premium rates to provide for varying rates of non-renewal (*J. I. A.*, Vol. XXXIII, 277, 278).

In obtaining the exposed to risk and rates of mortality and non-renewal we may in general record the data as shown in Chapter V by the policy year and nearest duration methods for a single decrement table, but we must separate the matured from the withdrawals. Furthermore we must include in $\sigma_{[x]+t}$ all policies terminated prior to the period of investigation at nearest duration t and those terminated at exact duration t . These additional policies will be included also in $w_{[x]+t}$ so that they will both enter and also withdraw at duration t . This will have no effect on the exposed to risk of death but will give the necessary data for the rate of non-renewal at duration t . If $\sigma_{[x]+t}$ were not to include these cases, such entrants would have to be excluded from the exposed to risk of non-renewal at duration t , making it necessary also to exclude the withdrawals at durations between t and $t + .5$ among such entrants. This could be done by excluding from $\sigma_{[x]+t}$ all cases terminated by withdrawal at durations less than $t + .5$ and half of the withdrawals at that exact point—as this would not affect $E_{[x]+t}$ —and by excluding $\sigma_{[x]+t}$ from the exposed to risk of non-renewal at duration t . Let us assume that the former course has been adopted in order to give additional data.

Similarly for the existing, let us suppose that $e_{[x]+t}$ excludes policies withdrawn at nearest duration t even though they may have withdrawn after the policy anniversary, and that the policies so excluded are included in $w_{[x]+t}$. Then $e_{[x]+t}$ will be included in the exposed to risk of non-renewal at duration t .

If we represent the matured at nearest duration t by $T_{[x]+t}$, formula (3) for the exposed to risk at death will become

$$E_{[x]+t+1} = E_{[x]+t} + \sigma_{[x]+t+1} - (T_{[x]+t+1} + w_{[x]+t+1} + e_{[x]+t+1} + \theta_{[x]+t}) \quad (18)$$

and

$$q_{[x]+t} = \frac{\theta_{[x]+t}}{E_{[x]+t}},$$

the rate of mortality and also the probability of death in the double decrement table. Also $(nrE)_{[x]+t}$, the exposed to risk of non-renewal at the end of the (t) th year,

$$\begin{aligned} = n_{[x]} + \sum_1^t \sigma_{[x]+t} - \sum_0^t T_{[x]+t} \\ - \sum_0^{t-1} w_{[x]+t} - \sum_1^{t-1} e_{[x]+t} - \sum_0^{t-1} \theta_{[x]+t}. \end{aligned}$$

Then

$$\begin{aligned} (nrE)_{[x]+t+1} &= (nrE)_{[x]+t} + \sigma_{[x]+t+1} \\ &\quad - (T_{[x]+t+1} + w_{[x]+t} + e_{[x]+t} + \theta_{[x]+t}) \\ &= E_{[x]+t} + \sigma_{[x]+t+1} - T_{[x]+t+1} - \theta_{[x]+t}. \end{aligned}$$

If then we let

$$f_{[x]+t} = w_{[x]+t} + e_{[x]+t}$$

and

$$g_{[x]+t} = \sigma_{[x]+t} - T_{[x]+t} - \theta_{[x]+t-1}$$

we have

$$E_{[x]+t+1} = E_{[x]+t} + g_{[x]+t+1} - f_{[x]+t+1} \quad (19)$$

and

$$(nrE)_{[x]+t+1} = E_{[x]+t} + g_{[x]+t+1}. \quad (20)$$

The rate of non-renewal at the end of the (t) th year is

$$(nrq)_{[x]+t} = \frac{w_{[x]+t}}{(nrE)_{[x]+t}}.$$

Having calculated the rates $q_{[x]+t}$ and $(nrq)_{[x]+t}$ we can construct the double decrement table by the relation

$$\begin{aligned} l'_{[x]+t+1} &= (l'_{[x]+t} - q_{[x]+t} l'_{[x]+t})[1 - (nrq)_{[x]+t+1}] \\ &= l'_{[x]+t}(1 - q_{[x]+t})[1 - (nrq)_{[x]+t+1}]. \end{aligned} \quad (21)$$

This form of the equation shows that the probabilities of living and of renewal are entirely independent of each other.

The data which are used as the basis of the rates of mortality and withdrawal may include policies issued with deferred dividend periods. Such policies were so often sold with emphasis on the total cash surrender value available at the end of the dividend period, that it may be considered advisable to treat them as matured on the dividend due-date.

Another point which may be mentioned is that it may be found, in any particular experience, that the non-renewal or withdrawal rate may vary with duration only for a certain number of years. The rate might even become practically constant for all attained ages beyond a fixed duration. In either case tables of mortality and withdrawal could be calculated in the select and ultimate form, the duration of the select period being the greater of the select period of mortality and the select period of withdrawal.

Tables of mortality and withdrawal have been the commonest forms of double decrement tables based on the records of insured lives in America, but in recent years tables having decrements from mortality and total and permanent disability have assumed great importance. The probabilities involved together with formulae for the exposed to risk are given in Actuarial Studies No. 5.

In Great Britain the subject of marriage and mortality has been of importance to actuaries because of such contingencies that have been occasionally insured against by the British Companies. Dr. Sprague's paper presented in 1879 (*J. I. A.*, Vol. XXI, 406) has already been mentioned. His formulae, developed on the assumption of a uniform distribution of deaths and marriages throughout the year, are applicable to that subject, though introducing a considerable error when used for withdrawals in ordinary insurance experience. In a later paper (*J. I. A.*, Vol. XXVIII, 350) Dr. Sprague has given tables for the re-marriage of widowers, showing the necessity for a select period for "recent widowers" merging into an ultimate, or "chronic widowers" rate at the end of twenty years for those who become widowers at age 25, the select period shortening as the age increases. In a paper in the same volume (p. 384) Mr. J. Chatham brings out that in constructing the select portion of such tables we may obtain $(wl)_{[x]+t}$, the widowers,

becoming such at age x , who are alive and have not re-married at age $x + t$, by working backward from $(wl)_{[x]+t+1}$ by the relation

$$(wl)_{[x]+t} = (wl)_{[x]+t+1} \cdot \frac{1 + \frac{1}{2}(\mu_1 + \mu_2)}{1 - \frac{1}{2}(\mu_1 + \mu_2)},$$

where μ_1 and μ_2 are respectively the central marriage and death rates at age $[x] + t$, and similarly for the individual decrements:

$$(wm)_{[x]+t}, \text{ widowers marrying, } = (wl)_{[x]+t+1} \cdot \frac{\mu_1}{1 - \frac{1}{2}(\mu_1 + \mu_2)}$$

$$(wd)_{[x]+t}, \text{ widowers dying, } = (wl)_{[x]+t+1} \cdot \frac{\mu_2}{1 - \frac{1}{2}(\mu_1 + \mu_2)}.$$

These equations are obtained by transforming the equations corresponding to those contained on p. 53.

Messrs. Hewat and Chatham investigated the mortality and marriage experience of the Widows' Funds of the Scottish Banks (*J. I. A.*, Vol. XXXI, 428) by calculating the central rates. The period of investigation was closed for the different banks in such a way as to lead to the assumption that the existing were exposed to risk on the average for six months after the age at which they were recorded as terminating. They describe their method as being analogous to the exact duration method. Using $(bm)_x$ to indicate the number of bachelors marrying between the ages x and $x + 1$ and $(bd)_x$ to indicate the bachelors dying within that year of age, they employed the following formula for the number of bachelors exposed to risk of both marriage and death at central age $x + \frac{1}{2}$.

$$\begin{aligned} E_{x+\frac{1}{2}} &= \Sigma(n_{x-1} - f_{x-1}) + n_x - \frac{1}{2}[e_x + w_x + (bm)_x + (bd)_x] \\ &= \Sigma(n_{x-1} - f_{x-1}) + n_x - \frac{1}{2}f_x, \end{aligned}$$

where

$$f_x = e_x + w_x + (bm)_x + (bd)_x.$$

It is important to note the simplicity of this method which requires no separation of both $(bm)_x$ and $(bd)_x$ into two parts similar to Mr. Ackland's method of treating withdrawals and deaths. In the above formula it is evident that the withdrawals, w_x , were recorded as terminating at their curtate durations, as well as those marrying, dying, and existing. The accuracy of assuming one half a year of exposure beyond curtate durations for all terminations should of course be tested before the assumption is made.

Mortality tables based upon the records of insured lives which contain more than two decremental factors would rarely be constructed. The principles involved are the same as those covered by the preceding discussion of double decrement tables. The computation of central rates will usually be found of advantage, because the exposed to risk has the same value for all central rates at the same age, and from the central rates the table may be constructed and also the annual rates (per year of exposure) may be calculated. These latter rates will often be wanted in order that they may be compared with similar rates from other experiences. It might also happen that some of the rates entering into the multiple decrement table are to be obtained from the investigation while others are to be taken from a different experience. In such a case we must obtain the central or annual rates, because we cannot compute the probability of a life surviving and otherwise remaining in the required status unless all the contingencies, as for example, death, total and permanent disability and withdrawal, entered into the single investigation.

CHAPTER VII. MISCELLANEOUS.

There will now be considered various points in connection with the construction of mortality tables which have not been previously discussed.

Final Series Method: In compiling the mortality experience of a life insurance company, it will be found that while there may be many old policies which have run their course, there will be a considerable number existing at the close of the observations. In the preparation of the mortality experience of the Thirty American Life Offices, which was based on calendar years and is described in Actuarial Studies No. 1, it was found that the average duration of the policies was only 4.36 years, and that of 982,734 male lives entering into the investigation 527,157 were existing at the close of the observations. The method of Final Series was adopted to carry the existing forward to their ultimate destination of death or withdrawal, and this was done on the assumption that the future experience on the existing would follow the same select rates of mortality and withdrawal that had actually been experienced in the past. Incorporating this hypothetical data obviously has an effect only on aggregate tables. The method employed may best be illustrated by an example taken from page 36 of the volume* giving the results of the investigation. The table on page 62 for one age at entry represents amounts to the nearest thousand with 000 omitted.

Although the values given represent thousands of dollars of insurance it will be easier to discuss them as though they represented lives or policies.

The original data show that there are 11,763 existing at the end of year 0 (*i.e.*, the first calendar year of insurance) to be carried forward, and out of 176,799 entrants of year 1 there are 27,361 discontinued and 1743 in claims in that year. In the final series the entrants on year 1 include the original 176,799 entrants and the 11,763 existing at the end of the previous year, making a total of 188,562. It is assumed that the discontinued and claims arising out of the 188,562 will be in proportion to the discontinued and claims arising out of the 176,799 entrants in

* System and Tables of Life Insurance—Meech, Vol. I.

CONSTRUCTION OF MORTALITY TABLES

Years of Insur- ance.	Experience Ending 1874.					Final Series.				
	Existing.	Discon- tinued.	Claims.	Entrants.	Entrants.	Discon- tinued.	Claims.	Exposed to Risk.	Multi- plier.	
0	11,763	7,443	457	196,462	196,462	7,443	457	94,509.5	1.0000	
1	10,058	27,361	1,743	176,799	188,562	29,181	1,859	173,971.5	1.0665	
2	8,433	15,564	1,250	137,637	157,522	17,813	1,431	148,615.5	1.1445	
3	7,692	10,655	959	112,390	138,278	13,109	1,180	131,723.5	1.2304	
4	8,705	7,753	916	93,084	123,989	10,327	1,220	118,825.5	1.3320	
	102,484	84,429	9,549							

the observed data. The discontinued and claims in the final series will therefore be found by multiplying the amount of these items in the original data by $188,562/176,799$, namely, by 1.0665. This gives 29,181 discontinued and 1,859 in claims, a total of 31,040; which subtracted from 188,562 gives 157,522, the entrants on year 2. It may be pointed out that the number of entrants on year 0 is the total of those passing out of observation as the existing, discontinued and claims, namely, $102,484 + 84,429 + 9,549$, or 196,462; and if the adjusted sum of discontinued and claims is deducted each year the balance gives the entrants on the following year. In this way the total existing according to the original data is automatically taken care of in the final series by merely calculating the amounts passing out of observation by discontinuance and claims.

The exposed to risk of death after year 0 is obtained by subtracting from the entrants one half the discontinued. For the year 0 the exposed is made up of half the entrants, less half the discontinued, or what is the same thing, one-half the difference between the entrants and discontinued for that year. The amount of claims divided by the exposed to risk of death gives the q_x values.

The exposed to risk and deaths for the first calendar year are unaffected; for the second year they are increased by a small percentage, for the third year by a larger percentage, and so on till the twenty-eighth year when the process terminated. The process as illustrated in the foregoing schedule resulted in a multiplier greater than 100 for some ages at the longest durations, a figure which appears to have alarmed the computers, as the multipliers finally adopted were those brought out by the original process multiplied by $(1.04)^{-n}$, where n represents the number of the year of insurance. Even with this adjustment the data of the later durations would at some ages be multiplied twenty and thirty fold. The hypothetical data for the aggregate table thus included an enormously larger proportion of experience at the longer durations than did the actual material, resulting in a considerable increase in the rate of mortality, especially at the older ages. It is open to question whether the final series method may not disturb the mortality rates to such an extent as to render the results untrustworthy. So far as we know it has been applied only in connection with the 30 American Offices' Experience.

King's Method: Mr. George King in Volume XXVII, page 218, of the *Journal* of the Institute of Actuaries, described a method of ascertaining the mortality experience of a life insurance company between valuation dates. It is quite customary for British life insurance companies to make a valuation of policy liabilities and to distribute surplus once in five years. A yearly valuation is sometimes made for the guidance of the companies, but the valuation on which the quinquennial distribution of surplus is based is made once in five years, and King's method was designed for the purpose of readily ascertaining the mortality experience of a company between valuation dates.

Assuming that the valuation date is December 31st, and changing the notation used by Mr. King, the method is as follows:

Let s_x = survivors at commencement of observations where x is the age nearest birthday on December 31st, the day before such commencement;

n_x = entrants during period of observation, where x is the age nearest birthday on nearest December 31st to date of entry;

w_x = withdrawals, where x is age nearest birthday on nearest December 31st to date of withdrawal;

θ_x = deaths, where x is age nearest birthday on December 31st preceding death;

e_x = existing at December 31st at close of observations, x being age nearest birthday at that date; and

E_x = exposed to risk in year of age x to $x + 1$.

Then:

$$E_x = E_{x-1} + s_x + n_x - w_x - \theta_{x-1} - e_x$$

and

$$q_x = \frac{\theta_x}{E_x}.$$

The s_x and e_x are taken directly from the classification registers which are used in connection with the valuation of the policy liabilities and it is necessary to classify merely the n_x , w_x and θ_x . Furthermore, fractional exposures are avoided. Although intended to apply only for a short term of years, this method might be extended to apply to the entire mortality experience of a company.

Mr. King points out that, as compared with the Institute Method, his formula avoids exposures of fractions of a year and is more accurate where entries and exits are not uniformly distributed throughout the year. Moreover the Institute Method gives no exposure to lives entering and discontinuing in the same calendar year, while approximate allowance is made for them under Mr. King's method. He calls attention also to the rapidity of the calculation of the exposed to risk by summation, as follows:

$$E_x = \Sigma(s_x + n_x - w_x - \theta_{x-1} - e_x)$$

and letting

$$s_x + n_x = h_x$$

and

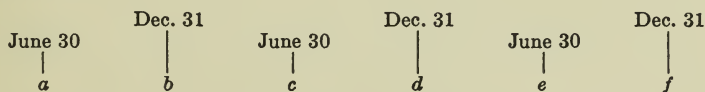
$$w_x + \theta_{x-1} + e_x = f_x;$$

then

$$\begin{aligned} E_x &= \Sigma h_x - \Sigma f_x \\ &= \Sigma(h_x - f_x). \end{aligned}$$

Mr. King's method was intended to be applied in British Life Offices where the method of valuation is different from that followed in this country. The majority of the business was tabulated according to the office age and was in convenient form to apply the method. If it were desired to make use of this plan in an office using the American form of valuation, the necessary particulars would require additional calculation.

Mr. Whittall has contended (*J. I. A.*, XXXI, 184) that Mr. King's method is one of mean ages. Looking at the question graphically,



an entrant in the period *bc* is relegated to *b* and his nearest birthday at point *b* must lie between *a* and *c*; an entrant in the period *cd* is relegated to *d* and his nearest birthday at the point *d* must lie between *c* and *e*. This is equivalent to taking mean ages with years running from June to June instead of December to December. This is apparent if it be assumed that the office years run from June 30th to June 30th. Then an entrant in the period *cd* would be relegated to *c* and his nearest birthday at that point would lie between *b* and *d*, that is, within the calendar year. Similarly an entrant in the latter half of

the financial year, *de*, would be relegated to *e* and at that point the nearest birthday would lie between *d* and *f*, that is, within the calendar year. In the case of an office with a financial year ending June 30th, the same results would be attained by adopting a system of mean ages, *i.e.*, subtracting year of birth from year of entry. The same principle applies to the survivors, withdrawals, deaths, and existing, allowing for the proper incidence of each.

Mr. King, in discussing Mr. Whittall's paper, disagreed that it is a method of mean ages, because that method, as ordinarily applied, permits as much as a year's error in age in any particular case, while his use of the nearest birthday on the nearest December 31st, never permitted an error on the assumed date of entry or exit of more than six months. They both agreed that it was difficult to classify the method as one purely of policy or calendar years. Mr. King, however, had previously stated (*J. I. A.*, Vol. XXIX, 178) that "the averages were so taken that they really got policy years. The average was so arranged as to make the policy years harmonize with the financial years of the company."

Compilation of Select Tables where Data have not been Traced on Policy Year Basis: A good example of the formation of select tables from a calendar year experience is the select tables formed by Dr. Sprague from the data of the Mortality Experience of the Institute of Actuaries (*J. I. A.*, XXI, p. 229). The data followed calendar years and the lives were taken at age next birthday at entry. It was assumed that the lives entered on the average in the middle of the calendar year and attained age next birthday at the end of the year. The lives were therefore aged $x - \frac{1}{2}$ at the date of entry in the middle of the year, x being age next birthday at entry.

The values of $\frac{1}{3}p_x$, $\frac{1}{4}p_x$, $\frac{2}{4}p_x$, etc., were first calculated for each age $(x - \frac{1}{2})$ at entry, the original exposed to risk being modified to give an exposure of 3 months to those withdrawing in the calendar year of entry. Then, in order to obtain sufficient data the experience was taken in groups of five ages. To avoid giving undue weight to any particular age, 50,000 entrants were taken for age $(x - 2\frac{1}{2})$, 100,000 for ages $(x - 1\frac{1}{2})$, $(x - \frac{1}{2})$, $(x + \frac{1}{2})$ and $(x + 1\frac{1}{2})$, and 50,000 for age $(x + 2\frac{1}{2})$, making 500,000 in each group. By taking 50,000 at the youngest and oldest age in each group an even 100,000 at each age was ob-

tained, since these extreme ages are each included in two groups. The values of $l_{[x-2\frac{1}{2}]+1\frac{1}{2}}$, $l_{[x-2\frac{1}{2}]+1\frac{1}{2}}$, $l_{[x-2\frac{1}{2}]+2\frac{1}{2}}$, etc., were next calculated, being based upon the 50,000 assumed for the value of $l_{[x-2\frac{1}{2}]}$. Similar values for each age at entry were calculated and a combination of the values for the same durations for ages at issue $(x - 2\frac{1}{2})$, $(x - 1\frac{1}{2})$, $(x - \frac{1}{2})$, $(x + \frac{1}{2})$, $(x + 1\frac{1}{2})$ and $(x + 2\frac{1}{2})$ were taken as the values for the average age x at entry, the value of $l_{[x]}$ being 500,000. The total deaths at the end of $\frac{1}{2}$, $1\frac{1}{2}$, $2\frac{1}{2}$, etc., years, were then found and by an interpolation formula the total deaths at the end of 1, 2, 3, etc., years were calculated. Then by taking the first differences of the results the numbers dying in each year of insurance could be obtained. Having obtained in this way the necessary data for quinquennial ages at entry, graphic graduation was used to adjust (a) the rates of mortality for the first insurance year, (b) the ratio of the rate of mortality of the second insurance year to that of the first, and (c) the corresponding ratio between the third and second insurance years. The values of $q_{[x]+n}$ for $n = 0$, $n = 1$ and $n = 2$ being thus available, and for $n > 4$ being taken as those of q_{x+n} by the $H^{M(5)}$ Table, values for $n = 3$ and $n = 4$ were supplied by interpolation. Interpolation was used also to supply the values for other than quinquennial entry ages. In some cases, for values of n less than 5, $H^{M(5)}$ values of q_{x+n} were substituted for those found as above described because q_{x+n} by the $H^{M(5)}$ table was less.

A corresponding method would have to be followed in order to obtain select tables from an experience based upon life years.

However, when it is possible to obtain sufficient data select tables should always be constructed by a policy year method and not by an indirect method such as that employed by Dr. Sprague.

Construction of Mortality Tables from Limited Data: It not infrequently happens that the data upon which a mortality table is based are not sufficient to give reliable, if indeed any, results at the young and old ages. Where the data are graduated by assuming some law of mortality, as described in Actuarial Studies No. 4, the values for the young and old ages may be supplied by assuming that the law holds throughout the table.

When a table is not so graduated it may be necessary to substitute for certain portions of the table which is being prepared, the data from some standard table which are believed to approximate to the mortality of the class of lives on which the new table

is based. The junction of the data of the standard table with that of the new table may be smoothed off by graduation or interpolation.

Where the data are very limited throughout, the actual deaths may be compared with the expected deaths according to some standard table, grouping a sufficient number of ages to secure enough data to give a general idea of the ratio of the actual mortality to that shown by the standard table at various ages. The ratios for intermediate ages may then be found by interpolation and the new table formed by applying the adjusted ratios to the data of the standard table.

Homogeneity of Data: In constructing a mortality table our purpose is usually not only to find out what rates of mortality have been experienced among the lives investigated but also to have a table which can be applied to estimate future events. The latter purpose is the chief one in practically all cases. If we are to be able to apply a table with any confidence we must know the character of the lives that contributed the experience and any special conditions to which their mortality was subject. It therefore becomes obvious that our material must be sufficiently homogeneous so that the results can be said to apply to a reasonably stable group of lives.

A little consideration, however, will show that when we speak of "homogeneous data," we are using a relative term. It is impossible to imagine a large group of lives which is not made up of many subgroups each of which is subject to slightly different mortality rates. For example in insurance experience, those lives which are insured under plans calling for the lowest premium rates will, generally speaking, show the highest rates of mortality. Where those insuring have a free choice as to whether their insurance shall be upon the participating or non-participating plan, it will generally be found that those selecting non-participating policies will show the higher mortality, though there is not the same reason for supposing that a company issuing only non-participating policies will experience a higher mortality than another company granting insurance only upon the participating plan. The relative mortality of male and female lives differs at various ages. The residents of different sections of the United States are subject to different rates of mortality. Even if we disregard outside influences we shall find degrees of difference among lives insured in one company at standard rates due to

physical or occupational causes. Then again mortality rates will be influenced by such external causes as agency methods, or, over different periods of time, the advances made in medical science and sanitation and in the science of selection.

Mr. G. F. Hardy* has given the following statement of the problem and its proper solution.

"The Actuary constructs tables not merely to show what has happened in the past, but to enable him to forecast the future, and as he requires these tables as a basis for financial operations, considerations are introduced which do not arise in the treatment of purely statistical tables. Whatever class of events the Actuary may have to deal with, will be subject to change with the lapse of time. That portion of the class he has been able to observe lies necessarily in the past; the conclusions he has derived from their study he proposes to extend to the future. He must therefore consider how far the observed characters of the class are changing or permanent, and must endeavour to distinguish between changes representing permanent tendencies and those due merely to temporary fluctuations. In the selection of data suitable for his purpose the Actuary will aim on the one hand at a sufficiently broad basis both in space and time to eliminate the effects of local and temporary fluctuations, and on the other hand he will aim at obtaining as far as possible a homogeneous group of data. These two aims are more or less in conflict, and he will lean to the one side or the other, according to the object he has in view. Where, for example, that object is to produce a table that may be adopted as a general standard by various institutions, often differing considerably as to their individual experience, he must aim at a correspondingly broad foundation. In these circumstances it will not generally be possible to obtain a really homogeneous experience. If it is a question of the mortality of assured lives, for instance, this will be found to be affected by endless individual variations, age, sex, duration of assurance, occupation, civil condition, class of assurance, character of the insuring office, etc., etc., and from such material approximately homogeneous data could only be obtained by cutting up the experience into comparatively small groups and thus sacrificing all generality. This can be avoided in practice by first excluding all extreme variations. The sexes will be separately treated, lives so impaired as to prospects of longevity by personal health, family history, occupation, or residence in unhealthy districts as to be "rated up" will be excluded, as also classes of assurance that may be supposed subject to rates of mortality differing from the average. When the data has thus been trimmed of the extreme variations, a body of experience will generally remain not greatly shrunken from its original dimensions and in which the discontinuous variations are sufficiently numerous and individually unimportant to render the data for practical purposes homogeneous. The rates of mortality, or withdrawal, can then be treated as functions of the two remaining variables of importance, the age and the time elapsed from date of entry; or as functions of the age only from the point at which the factor of duration may be found to be unimportant."

* "The Theory of the Construction of Tables of Mortality, etc.," p. 16.

Spurious Selection: After a reasonably homogeneous set of lives have been taken to form select and ultimate tables there is still some question whether the effect of the selection of risks can be measured precisely by the difference between the ultimate rates and the select rates for the same attained age but for durations within the select period; nor can we be altogether sure that the effect of selection is felt for the exact period shown by the tables. In the first place it will be noted that the practical necessities of graduation may lead to the choice of a uniform period of selection at all ages at entry contrary to the evidence of the unadjusted figures, and may also result in an appreciable change in the select rates for the later part of the period at some ages in order to form a smooth junction with the ultimate rates.

Mr. W. P. Elderton introduced the term "Spurious Selection" (*J. I. A.*, Vol. XL, 221) to describe "the selection indicated by a difference in q which has arisen entirely from statistical processes." He examined the case where a table is formed by amalgamating two classes of data to which he applied the symbols $E_{[x]+t}$ and $E'_{[x]+t}$, the prime indicating the data subject to the heavier mortality rate. Using q_{x+t} and q'_{x+t} to indicate the respective ultimate rates of mortality to which the classes are subject, the rate of mortality that would be found for the $(t+1)$ th year from the amalgamated data would be

$$\frac{E_{[x]+t}q_{x+t} + E'_{[x]+t}q'_{x+t}}{E_{[x]+t} + E'_{[x]+t}}.$$

Similarly the rate for the same attained age arising from the next year of duration is

$$\frac{E_{[x-1]+t+1}q_{x+t} + E'_{[x-1]+t+1}q'_{x+t}}{E_{[x-1]+t+1} + E'_{[x-1]+t+1}}.$$

The rates q_{x+t} and q'_{x+t} are used in both cases since each class of data is assumed to have passed beyond its period of selection. It is seen, however, that if q and q' are unequal the resulting mortality rates for the amalgamated table will be unequal though they apply to the same attained age. If the latter expression is the greater then selection will appear to be still effective, whereas we know it has really ceased, but if the former is the greater we shall get the opposite effect to that of selection, or what may be called a tendency to conceal the effects of selection.

The analysis may be carried further, for

$$\frac{E_{[x]+t}q + E'_{[x]+t}q'}{E_{[x]+t} + E'_{[x]+t}} \geq \frac{E_{[x-1]+t+1}q + E'_{[x-1]+t+1}q'}{E_{[x-1]+t+1} + E'_{[x-1]+t+1}}$$

as

$$E_{[x-1]+t+1}E'_{[x]+t}q' + E_{[x]+t}E'_{[x-1]+t+1}q \geq E'_{[x-1]+t+1}E_{[x]+t}q' + E_{[x-1]+t+1}E'_{[x]+t}q$$

or as

$$E_{[x-1]+t+1}E'_{[x]+t}(q' - q) \geq E_{[x]+t}E'_{[x-1]+t+1}(q' - q)$$

or as

$$\frac{E'_{[x]+t}}{E_{[x]+t}} \geq \frac{E'_{[x-1]+t+1}}{E_{[x-1]+t+1}}.$$

In other words we shall get the appearance of selection that does not exist if the right-hand member of the last equation is the greater, *i.e.*, if the proportion of data subject to the heavier mortality is increasing with duration for the same attained age. We may come to this conclusion also from general reasoning, for in such a case the ultimate data will contain an undue amount of the heavier mortality data while the very early years of duration will be largely made up of the lower mortality data, so that in order to have the mortality rates pass from the low initial select rates to the heavy ultimate rates will require us to pass through an increased number of years of duration. Thus the apparent effect of selection is increased both by increasing the differences between the select and ultimate rates and also by increasing the period of selection.

If, on the other hand, the left-hand member of this equation is the greater, we obtain the contrary effect, our initial select rates being increased by the heavier mortality and the ultimate rates being lowered by the larger proportion of lower mortality, thus tending to offset the true effects of selection.

The former condition would seem to be the more likely one in a life insurance experience because of the fact that, generally, there has been evidence that mortality has decreased with advancing calendar year of entry. The longer the duration the larger would be the proportion of data arising from early issue, if the investigation is terminated for all policies on a certain date or in a certain year.

Mr. P. C. H. Papps investigated another special case (*T. A. S. A.*, Vol. XIII, 211) in which spurious selection may be due to a progressive improvement in mortality with advancing calendar

years of exposure affecting all the data irrespective of year of issue. Such an effect might be due to improvement in the general health of the population. He assumed further that the data included only policies issued during the period of investigation. He has shown that, for an equal volume of issues each year, the result would be to produce too low an effect of selection.

This result could also be reached by general reasoning for the mortality rate for the longest year of duration would be determined from exposures in the last year of the investigation, while the rate for the first policy year would come from exposures over the entire period, with intermediate durations passing from one extreme to the other. Consequently the mortality rates for the early years of duration would be relatively high compared with those of the later, or ultimate years.

As pointed out by Mr. F. H. Johnston, in discussing Mr. Papps's paper, this effect vanishes when a table is constructed from data that includes the experience of issues prior to the period of investigation, assuming that the distribution of the exposed to risk for any one policy year was chronologically the same as for any other policy year—which is fairly close to the fact. In such a case the proportions of the exposed to risk at duration t arising in the first, second, third, etc., years of the investigation is the same as the proportions for other durations.* Then the mortality rate for each year of duration would be equal to a constant multiplied by the rate for that duration for any specified year of the investigation. This would also be true if the mortality rates changed in any manner from year to year of the investigation.

The question whether spurious selection exists is a complex one in any case, and its importance in any table constructed from carefully selected data in the usual way is probably small. It does emphasize, however, the need for obtaining homogeneous data if we are to place much reliance on the effect of selection displayed by the resulting tables.

* This would not be true in general for the longest durations, but this would have little effect on the ultimate rates in most cases.

UNIVERSITY OF ILLINOIS-URBANA

368.01AC8A

C001

ACTUARIAL STUDIES NEW YORK

2 1922



3 0112 016690056